2014-04-24

An Integrated Structural Mechanism for Relief of Autoinhibition and Membrane Targeting in Cytohesin Family Guanine Nucleotide Exchange Factors: A Dissertation

Andrew W. Malaby
University of Massachusetts Medical School

AN INTEGRATED STRUCTURAL MECHANISM FOR RELIEF OF AUTOINHIBITION AND MEMBRANE TARGETING IN CYTOHESIN FAMILY GUANINE NUCLEOTIDE EXCHANGE FACTORS

A Dissertation Presented

By

ANDREW W MALABY

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

APRIL 24, 2014

BIOCHEMISTRY AND MOLECULAR PHARMACOLOGY

AN INTEGRATED STRUCTURAL MECHANISM FOR RELIEF OF AUTOINHIBITION AND MEMBRANE TARGETING IN CYTOHESIN FAMILY GUANINE NUCLEOTIDE EXCHANGE FACTORS

A Dissertation by

ANDREW WILLIAM MALABY

Signature of the Dissertation Examination Committee
signifies completion and approval of style and content

_____

David Lambright, Ph.D., Thesis Advisor

_____

Osman Bilsel, Ph.D., Member of Committee

_____

J. Christopher Fromme, Member of Committee

_____

Lawrence Stern, Ph.D., Member of Committee

Signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee.

_____

William Royer, Ph.D., Chair of Committee

Signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school.

_____

Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Biochemistry and Molecular Pharmacology
April 24, 2014

To my parents: who rose the bar so I could jump, climb, swing, flail, struggle, flip, and fly over it. You taught me that through hard work and perseverance, I could find my place and earn my own perfect 10.

To my wife Heidi: my perfect 10.

ACKNOWLEDGMENTS

Many people have made this journey possible, and I could write another thesis describing their kind help along the way. First and foremost, I want to thank my lab. My mentor, David Lambright has helped me see the big picture in every experiment. I can't count the number of times I have walked into his office feeling confused, and out feeling motivated, or least confused about new things. Additionally, labmates Claudia Del Campo, Jianjun Duan, Gabe (Meng-Tse) Lee, and Ashwini Mishra have been great personal friends and professional colleagues.

A big thanks goes to the Biochemistry and Molecular Pharmacology and Molecular Medicine departments. It has been a pleasure to work in such a knowledgeable, diverse, and supportive environment that never takes itself too seriously. Special thanks to Silvia Corvera, Robert Matthews, Mary Munson, William Royer, and Lawrence Stern for all of the advice and support on Qualifying and Thesis Committees.

A great deal of gratitude is owed to my friends outside the lab. To my friends at Sterling Academy of Gymnastics and throughout the New England gymnastics community, Fellowship Church, Worcester Scrabble Club, and Seven Hills Symphony - you've been my diversion, distraction, and relief.

Thanks to my family for always supporting me, even when they didn't get it, especially in my first couple of years of this process when I needed it the most. My wife Heidi continues to share the highs and lows of research with me – the refrigerator gels and the voodoo notebooks. She is a constant source of strength in my life.

ABSTRACT

Guanine nucleotide exchange factors (GEFs) regulate and organize diverse cellular processes through their role in converting GTPases from the inactive GDP bound state to the active GTP bound state. An increasing number of GEFs undergo autoregulatory mechanisms through complex intramolecular interactions. Relief of autoinhibition involves specific phosphorylation or binding to lipid and/or effector proteins at sites distal from the catalytic domain, and is often coupled to membrane recruitment. In Cytohesin Arf GEFs, the catalytic Sec7 domain is autoinhibited by a linker region and C-terminal helix flanking a Pleckstrin Homology (PH) domain. Upon binding of the PH domain to low abundance phosphoinositides, the GTPase Arf6-GTP can both relieve autoinhibition and recruit Cytohesins to the plasma membrane. This thesis focuses on determining the molecular mechanism underlying both these functions.

The structural mechanisms by which Arf6-GTP binding relieves autoinhibition were studied using biochemical and crystallographic studies. The crystal structure of the Grp1 PH domain in complex with Arf6 revealed that Arf6-GTP binding relieves autoinhibition through competitive sequestration of the inhibitory elements into grooves formed at the periphery of the interface. Importantly, the interaction orients all known membrane targeting components to a common surface. Detailed biochemical studies showed a common mode of binding among Cytohesin family members in which phosphoinositide head group binding primes the interaction with Arf6, and membrane recruitment of both stimulatory and substrate Arf enhances the effect.

To assess changes in the Sec7 domain conformation upon activation, Size Exclusion Chromatography in line with Small Angle X-Ray Scattering (SEC-SAXS) was performed. The unique nature of this data led to the development of a novel data analysis and processing strategy. A graphically based, python-extensible software package was created for data normalization, buffer correction, Guinier Analysis, and constant background subtraction. As an unbiased substitute for traditional buffer subtraction, a method to reconstruct the protein scattering through singular value decomposition (SVD) and linear combination of the basis vectors was developed. These methods produced exceptional data quality and allowed versatility for application to other data collection techniques or systems, especially those lacking confident buffer matching or low signal.

SEC-SAXS confirmed the overall structure of autoinhibited Grp1 in solution and showed only slight overall changes upon activation by deletion of the autoinhibitory C-terminal helix. Fusion of Arf6 with Grp1 produced a consistently elongated shape in the active state that was incompatible with the autoinhibited or theoretical active positions of the Sec7 domain. Monte Carlo and rigid body modeling using known structural domains revealed a requirement for Sec7-PH linker flexibility in addition to Sec7 domain mobility. These data support an integrated structural model whereby phosphoinositides and Arf-GTP support nucleotide exchange at membranes through allosteric activation, membrane recruitment, and large-scale rearrangement of the Sec7 domain. Overall, these findings offer insight into Cytohesin function that can be applied to assess relief of autoinhibition in the context of other GEFs and GTPases.

TABLE OF CONTENTS

TITLE PAGE i

SIGNATURE PAGE ii

DEDICATION iii

ACKNOWLEDGEMENTS iv

ABSTRACT v

TABLE OF CONTENTS vii

LIST OF TABLES viiii

LIST OF FIGURES x

ABBREVIATIONS xii

PREFACE xiv

CHAPTER I: INTRODUCTION 1

Abstract 1
Membrane Trafficking and Disease 2
GTPases as Molecular Switches and Organizers of Cellular Trafficking 4
Arf GTPases 10
Arf GEFs and GAPs 20
Cytohesin GEFs 33
Autoregulatory Interactions in Guanine Nucleotide Exchange Factors 36
Thesis Rationale 44

CHAPTER II: STRUCTURAL BASIS FOR MEMBRANE RECRUITMENT AND ALLOSTERIC ACTIVATION IN CYTOHESIN FAMILY ARF GTPASE EXCHANGE FACTORS 46

Abstract 46
Introduction 47
Results 48
Discussion 74

Materials and Methods 79

CHAPTER III: DEVELOPMENT OF DATA PROCESSING AND ANALYSIS TECHNIQUES FOR SIZE EXCLUSION CHROMATOGRAPHY SMALL ANGLE X-RAY SCATTERING 85

Abstract 85
Introduction 86
Results 90
Discussion 111
Materials and Methods 114

CHAPTER IV: STRUCTURAL ARCHITECTURE AND DYNAMICS IN ACTIVATION OF CYTOHESIN FAMILY GUANINE NUCLEOTIDE EXCHANGE FACTORS 119

Abstract 119
Introduction 121
Results 123
Discussion 145
Materials and Methods 150

CHAPTER V: DISCUSSION 154

Unified Mechanisms for Autoinhibition Relief in Cytohesin Family GEFs 154
Relief of Autoinhibition by Arl4d 162
Mechanisms for Relief of Autoinhibition in the Pathogenic GEF RalF 163
Overall Implications and Concluding Remarks 168
Materials and Methods 169

LIST OF TABLES

Table II.1: Data collection and refinement statistics for the Grp1-IP_4:Arf6-GTP Crystal Structure
Table II.2: Catalytic efficiencies for $Grp1_{63-399}$ with active Arf GTPases
Table II.3: Catalytic efficiencies of Cytohesin variants

Table IV.1: Radius of Gyration and D_{max} values from SEC-SAXS.

LIST OF FIGURES

Figure I.1: Cellular Trafficking
Figure I.2: The GTPase Cycle
Figure I.3: GTPase Networks
Figure I.4: Cellular Localization of Arf Family GTPases
Figure I.5: Domain Architecture and Sequence Conservation of Arf Family GTPases
Figure I.6: Arf GTPase Structural Elements
Figure I.7: Arf:Effector Specificity Determinants
Figure I.8: The ARNO Sec7 Domain in Complex with Arf-GDP Substrate
Figure I.9: Overall Structure of Representative PH Domains
Figure I.10: Phosphoinositide Recognition in PH Domains
Figure I.11: Variable Phosphoinositide Binding Geometries in PH domains
Figure I.12: Autoregulation in Sos and Vav1
Figure I.13: Autoregulation in Sec7 and BRAG2
Figure I.14: The Structure of Autoinhibited Grp1

Figure II.1: Relief of Autoinhibition by Arf6-GTP
Figure II.2: Direct Binding of Arf6-GTP to Grp1
Figure II.3: The Grp1-IP_4:Arf6-GTP Crystal Structure
Figure II.4: The Grp1 PH Domain Binding Epitope on Arf6-GTP
Figure II.5: Interactions of the Grp1 Autoinhibitory Elements within the Grp1-Arf6 Binding Interface
Figure II.6: Detailed Interaction Map of the Grp1-IP_4:Arf6-GTP Complex
Figure II.7: Competitive Allosteric Mechanism for Relief of Autoinhibition
Figure II.8: Mutational Analysis of the Grp1-IP_4:Arf6-GTP Interaction
Figure II.9: Effect of Cytohesin Splice Variants on Arf6-GTP Binding and Nucleotide Exchange
Figure II.10: Experimental Setup for Studying Arf6 Stimulation of Cytohesin Activity in Membrane Environments Using Soluble Substrate
Figure II.11: Atf6-Stimulation of Cytohesin Activity in Membrane Environments Using Soluble Substrate
Figure II.12: Arf6-Stimulation of Cytohesin Activity in Membrane Environments Using Membrane Associated Substrate
Figure II.13: Model for Integrated Relief of Autoinhibition and Membrane Recruitment of Cytohesins by Arf6-GTP

Figure III.1: SEC-SAXS Analysis for Arf6 and Grp1
Figure III.2: Data Normalization of $Grp1_{63-399}$
Figure III.3: Buffer Subtraction Trials for Arf6 and $Grp1_{63-399}$
Figure III.4: Singular Value Decomposition (SVD) of $Grp1_{63-39}$
Figure III.5: Linear Combination of SVD Components for $Grp1_{63-399}$ and Arf6

Figure III.6: Characterization of the $Grp1_{63-399}$ dimer species by SEC-SAXS
Figure III.7: Constant Subtraction Scouting by Porod Plot
Figure III.8: Effects of Constant Subtraction on *Ab Initio* Shape Envelopes
Figure III.9. The $Grp1_{63-399}$ Crystallographic Dimer

Figure IV.1: SEC-SAXS analysis of $Grp1_{63-399}$ and $Grp1_{63-390}$.
Figure IV.2: Domain Structure and Kinetic Analysis of Grp1-Arf6 Tandem Constructs
Figure IV.3: SEC-SAXS Analysis of Select Grp1-Arf6 Tandem Constructs
Figure IV.4: Domain Assignment of Grp1-Arf6 Tandem Shape Envelopes
Figure IV.5: Structural Rearrangement in Active Grp1-Arf6 Tandem Constructs
Figure IV.6: Ensemble Optimization Method Analysis of Grp1-Arf6 Tandems with Flexible Linkers
Figure IV.7: Rigid Body Modeling of Ensemble Optimization Method Structures for Grp1-Arf6 Tandems
Figure IV.8: Orientations of the Grp1 Sec7 Domain after Rigid Body Refinement with Flexible Sec7-PH Domain Linker
Figure IV.9: Model for Relief of Autoinhibition with Sec7 Domain Reorientation

Figure V.1: SAXS Total Intensity Chromatograms for Dimeric Grp1 Constructs
Figure V.2: Analysis of Grp1 Dimers by SEC-SAXS
Figure V.3: Model for Autoinhibition Relief in Dimeric Cytohesins
Figure V.4: Stimulation of Grp1 Activity by Arl4d-GppNHp
Figure V.5: Liposome Partitioning and Catalytic Activity of RalF

LIST OF SYMBOLS AND ABBREVIATIONS

2G	Diglycine
3G	Triglycine
Å	Ångström
app	Apparent
Arf	ADP Ribosylation Factor
ARHGAP	Rho GTPase Activating Protein
Arl	Arf-like
ARNO	ADP Ribosylation Factor Nucleotide Site Opener
BRAG1	IQSEC1: IQ and Sec7 domain
CASP	Cytohesin-Associated Scaffolding Protein
CtH	C-terminal Helix
CYTH1	Cytohesin-1
CYTH4	Cytohesin-4
DAPP	Dual Adaptor for Phosphoinositide and 3-phosphoinositides
DH	Dbl Homology
EDTA	Ethylenediaminetetraacetic Acid
EOM	Ensemble Optimization Method
ER	Endoplasmic Reticulum
FAPP1	Four Phosphate Adaptor Protein
FERM	4.1 protein/Ezrin/Radixin/Merlin Homology
GAP	GTPase Activating Protein
GDP	Guanine Diphosphate
GEF	Guanine Nucleotide Exchange Factor
GRASP	Grp1-Associated Scaffolding Protein
Grp1	General Receptor for Phosphoinositides
Grsp1	Grp1 Signaling Partner
GS	Glycine-Serine
GDP	Guanosine Diphosphate
GTP	Guanosine Triphosphate
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HR	Heptad Repeat
Ins(1,3,4,5)P_4	Inositol-(1,43,4,5)-tetrakisphosphate
ISW	Interswitch
JIP4	JNK-Interacting Protein 4
$K_{0.5}$	Half Maximal Activity
k_{cat}/K_M	Catalytic Efficiency
KD	Dissociation Constant
LUV	Large Unilamellar Vesicles
mant	Methylanthraniloyl
MBD	Membrane Binding Domain
myr	Myristoylated
P(r), PDDF	Pairwise Distribution

PBR	Polybasic Region
PH	Pleckstrin Homology
PI	Phosphatidyl Inositol
PI3K	Phosphinositol-3-kinase
PKC	Protein Kinase C
PLC1-δ	Phospholipase C delta
PM	Plasma Membrane
PMSF	Phenylmethanesulfanylfluoride
POPC	Palmitoyloleolylphosphatidylcholine
POPS	Palmitoyloleolylphosphatidylserine
PtdIns	Phosphatidylinositol
Sar	Secretion-Association Ras-related GTPase
SAXS	Small Angle X-ray Scattering
Sec	Secretion
SEC	Size Exclusion Chromatography
Sos	Son of Sevenless
SPR	Surface Plasmon Resonance
SUMO	Small Ubiquitin-Like Modifier
SVD	Singular Value Decomposition
SW1	Switch I
SWII	Switch II
TAPP	Tandem PH Domain Containing Protein
Tris	Tris(hydroxymethy)aminomethane

PREFACE

CHAPTER I is a general Introduction for the dissertation. Andrew Malaby generated all figures.

CHAPTER II is adapted from a published manuscript:

> Malaby AW, van den Berg B, Lambright DG "Structural basis for membrane recruitment and allosteric activation of cytohesin family Arf GTPase guanine nucleotide exchange factors." *Proceedings of the National Academy of Sciences, USA*. 2013 Aug 27;110(35):14213-8
>
> Andrew Malaby did all biochemistry, crystallization, and structural refinement described and Dr. Bert van den Berg collected synchrotron diffraction data. Dr. David Lambright helped with design of the membrane targeting model.

CHAPTERS III and IV are in preparation for a unified publication. For clarity, they are expanded into two chapters covering single areas of method development and analysis of the biological system. The development of the software package DELA and its use for SAXS data processing is described in part in:

> Lambright DG, Malaby, AW, Kathuria SV; Nobrega RP, Bilsel O, Matthews CR, Muthurajan U, Luger K, Chopra R, Irving TC, Chakravarthy S "Complementary techniques enhance the quality and scope of information obtained from SAXS." (*Trans Amer Crys Soc.* 44:1-12)
>
> Andrew Malaby carried out all biochemistry, SAXS, and Monte Carlo and rigid body modeling experiments. Dr. David Lambright wrote and developed the software package and scripts for data processing and analysis. Both were involved in method development.

CHAPTER V includes preliminary data for future studies related to studying autoinhibition in dimeric cytohesins and with other GEFs and stimulatory GTPases.

> Andrew Malaby did all experiments and analysis. Plasmids encoding RalF were donated from the lab of Dr. Craig Roy. Dr. Jonathan DiNitto generated plasmids encoding $Grp1_{13\text{-}399}$, $Grp1_{13\text{-}390}$, and $Arno_{2\text{-}400}$ and purified the proteins.

CHAPTER I: INTRODUCTION

Abstract

Membrane trafficking is an essential process in eukaryotes. Throughout the cell, GTPases act as molecular switches and organizers of these events through their ability to couple membrane localization and recruitment of effector molecules with a nucleotide cycling mechanism. Nucleotide binding and thus effector function are elegantly regulated through structural mechanisms that allow selectivity of diverse effectors for specific nucleotide states. One key example of GTPases is the Arf family, which controls membrane traffic at both internal organelles and the cell periphery. These and other GTPases require diverse effector molecules to facilitate nucleotide exchange and hydrolysis.

Sec7 family guanine nucleotide exchange factors (GEFs) and Arf GAP family GTPase activating proteins (GAPs) control nucleotide cycling process through well studied mechanisms of GDP-GTP exchange and GTP hydrolysis. In addition to catalytic domains, many GEFs and GAPs have membrane binding and regulatory domains to organize events at particular cellular locations. Membrane binding domains such as the PH domain have evolved for specialized recognition of phosphoinositides that incorporates binding of diverse species through variable geometries. Autoregulatory interactions serve to modulate catalytic activity based on specific protein or membrane interactions. Autoregulation provides key integration of GTPase, effector, and membrane inputs and may represent a unified theme in the organization of complex trafficking networks into more discrete pathways for cellular function.

Membrane trafficking and disease

The packaging and traffic of cellular components between various organelles is vital to cellular homeostasis and function. Channel components functioning at the cell periphery, for example, need to undergo trafficking from the endoplasmic reticulum (ER) to the plasma membrane (PM). Similarly, protein receptor levels at the PM are continually adjusted by endocytosis of receptors and subsequent recycling or degradation (Figure I.1).

Due to the complexity of cellular transport, a wide range of disease states occur when trafficking is compromised by defects in the cargo molecules, defects in the trafficking molecules, or external infection [1-4]. For example, mutation of the insulin receptor gene results in decreased trafficking of receptors to the cell surface and therefore widespread insulin resistance associated with diabetes [5]. Mutations in the dynein motors which create movement on scaffold microtubules can cause developmental defects in the lungs and neurons [6]. Additionally, bacterial pathogens such as *Legionella pneumophila* hijack host trafficking pathways by injecting competing trafficking proteins into the cell, avoiding degradation by the cell's defense mechanisms and fostering widespread proliferation of the infection [7, 8].

The wide spectrum of disorders from defects in transport is evidence that trafficking is a complex process for homeostatic regulation and specialized function as might be the case upon hormone stimulation. Myriads of on-off interactions at all membrane bound organelles, each a unique milieu of diverse membrane compositions and internal cargo, have to be coordinated for efficient, dynamic transport of materials.

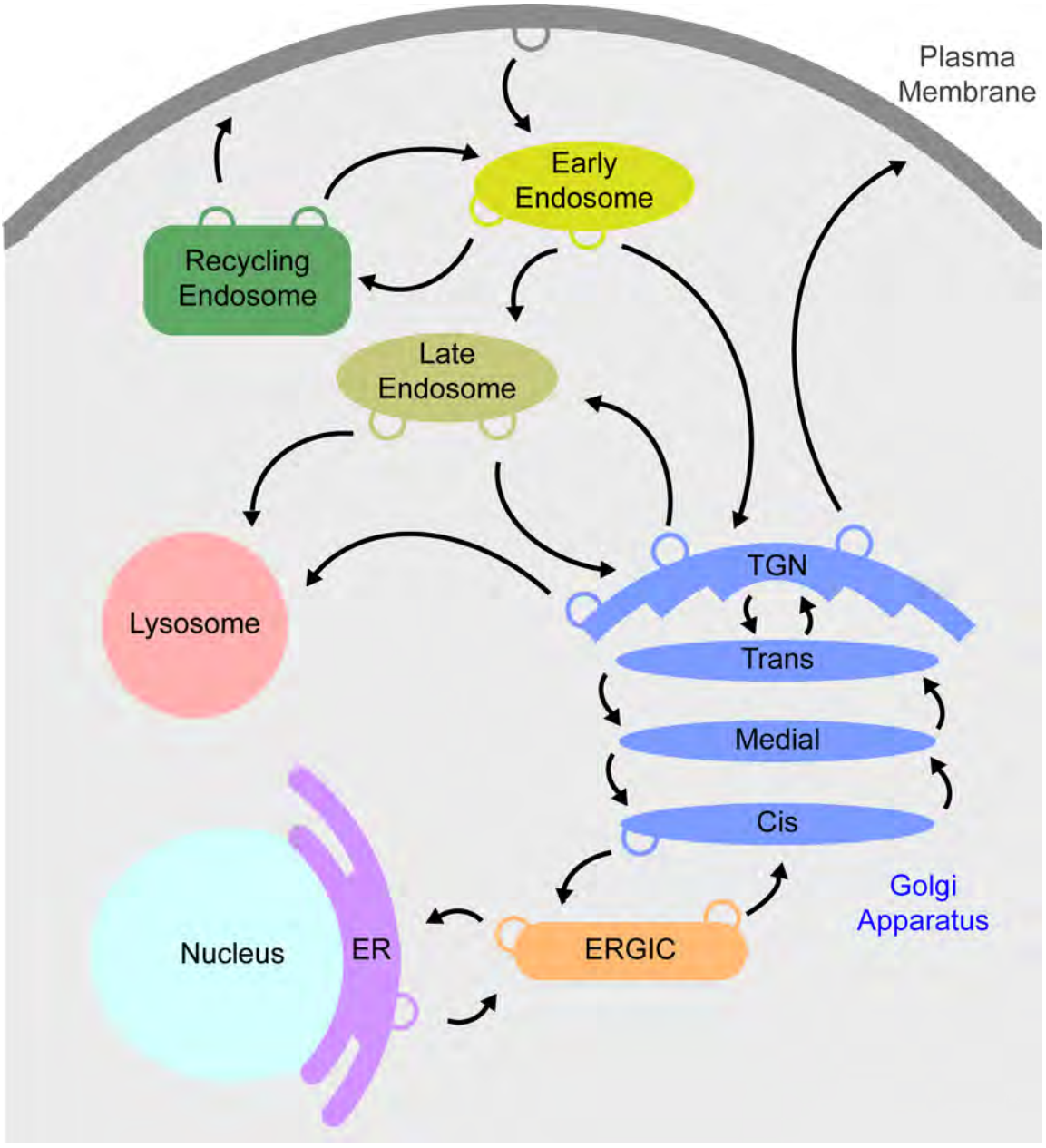

Figure I.1: Cellular Trafficking. Schematic representing steps of endocytosis and exocytosis. ER, Endoplasmic Reticulum; TGN, Trans Golgi Network; ERGIC, ER-Golgi Intermediate Compartment

Spatiotemporal organization comes from an incredibly large array of proteins with versatile, but also regulated function at every key step in the process

GTPases as Molecular Switches and Organizers of Cellular Trafficking

The GTPase Cycle

Membrane trafficking requires versatile but regulated effector binding to effect widespread changes. One group of molecules serving this function is small GTPases, which serve as molecular switches by binding to either guanosine diphosphate (GDP) or guanosine triphosphate (GTP). The GDP-bound state is considered inactive, whereas the GTP-bound state is the "active" state capable of binding classical effectors to promote cellular trafficking events (Figure I.2A). Besides controlling effector binding, nucleotide state also plays a role in membrane recruitment of some GTPases [9, 10].

The nucleotide state of GTPases is actively regulated. To overcome slow intrinsic kinetics for nucleotide exchange, the cell uses guanine nucleotide exchange factors (GEFs) and GTPase Activating Proteins (GAPs). GEFs activate GTPases by facilitating GDP release, allowing GTP to bind. GAPs increase the intrinsic levels of hydrolysis of GTP to GDP on the protein [11]. Further regulation comes from GDP or GTP-specific effectors that can sequester or recruit GTPases as well as complex regulation of GEF and GAP function which will be focused on in subsequent sections [12, 13].

Membrane Recruitment of GTPases

In addition to their roles as molecular switches, GTPases facilitate cellular trafficking through transient recruitment to membrane environments via post-translational addition of fatty acid and/or prenyl groups. Although the nature and location of the specific modifications varies greatly, each modification cooperates with other factors such as membrane insertion elements to provide stable membrane attachment [14]. For example, the protooncogenic kinase Src is stabilized at membrane environments through a polybasic region as well as a myristate group at the N-terminus [15-19], and the GTPases N-and H-Ras use both C-terminal farensylation as well as palmitoylation for membrane docking [20, 21].

Fatty acid or prenyl attachments to the membrane can be controlled in multiple ways. Palmitoylation is controlled by expression, localization, and activity of palmitoyl acyltransferases (PATs) and depalmitoylases. In the case of Ras GTPases, expression of PATs in the Golgi and depalmitoylases at the PM membrane coordinate regulation of the GTPase distribution in the cell. Prenyl modifications can be stably associated with membranes until extraction by effector molecules. A notable example of this is membrane extraction of Rab and Rho GTPases by GDI (guanine nucleotide dissociation inhibitor) through sequestration of geranylgeranyl groups on the GTPase [14]. Although Rab GEFs are essential and sufficient for membrane targeting upon activation in many cases, in other cases, other Rab effectors contribute, such as GDF (GDI displacement factor), which competitively binds GDI [22].

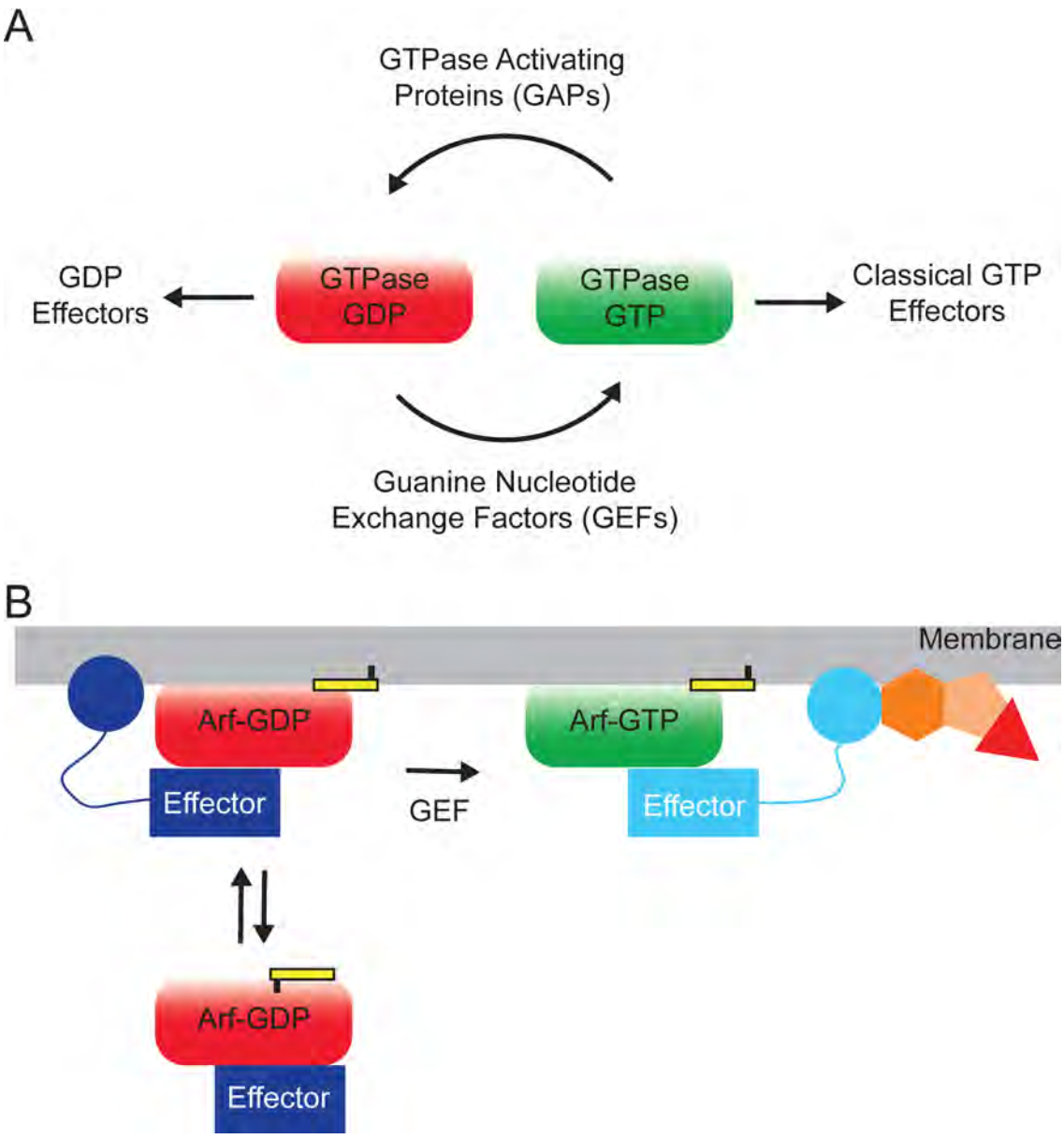

Figure I.2: The GTPase Cycle. A) Schematic of the GTPase 'molecular switch' in which GDP or GTP bound states control effector binding. B) Schematic of Arf GTPases as organizers of membrane traffic. Arf GTPases can recruit primary and secondary effectors onto membrane environments, localizing trafficking signals to sites of GTPase activation.

The use of membrane anchoring switches is a key way GTPases can use nucleotide cycling to act as organizers of cellular transport. This is exemplified by the myristoyl switch in Arf GTPases (Figure I.2B). Unlike other fatty acid modifications, myristoyl groups are permanently attached to an amphipathic helix via N-myristoyl transferases, as no mammalian demyristoylase has been identified to date [23]. To regulate membrane association, the helix is sequestered in the GDP bound state and exposed when bound to GTP, promoting insertion of the hydrophobic group into the membrane [24]. Association with membranes allows organization of many factors at sites of action.

GTPase Networks.

In addition to transducing signals through membrane attachments at specific locations, GTPases often are involved in coordination, amplification, and crosstalk among or between trafficking steps in GTPase networks, first clearly demonstrated in Rab GTPases (Figure I.3) [12]. Coordination of multiple steps can be achieved by a GTPase recruiting the GEF or GAP for a downstream GTPase. An example is in exocytosis from the Golgi, in which subsequent secretion in yeast is coordinated by the GTPase Ypt31/32 and continued on secretory vesicles by another GTPase Sec4, a Rab8 homolog. Ypt32 coordinates these processes by recruiting Sec2, the GEF for Sec4 [25]. Importantly, Sec2 is not a GEF for Ypt32, confirming coordination rather than competition of these GTPases. Similarly, in endocytosis from early to late endosomes, the early endosome GTPase Rab22 recruits the Rab5 GEF Rabex-5, which in turn can

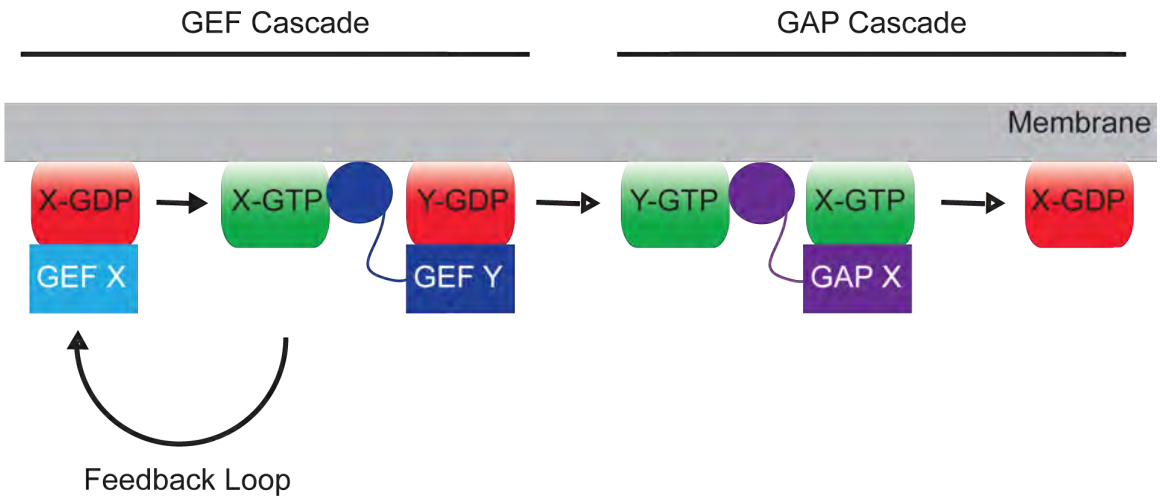

Figure I.3: GTPase networks. Examples of GEF or GAP cascades in which the GEF or GAP for one GTPase recruits a GEF or GAP for the sequential GTPase. Left) Feedback loop in which an active GTPase stimulates it's own GEF..

recruit the GEF complex for the late endosome GTPase Rab7 [26-28]. These “Rab cascades” for GEFs and GAPs have since been extended to other GTPase families at many cellular compartments [12].

In select cases, feedback loops have emerged to amplify or terminate signals within GTPase networks. These loops highlight GTPases as central directors of widespread cellular processes rather than simple checkpoints for specialized downstream functions [29]. The oncogenic Ras GTPase and its GEF Son of Sevenless (Sos) typify this phenomenon. Ras activation promotes recruitment and activity of Sos to specific plasma membrane environments [30]. As will be expanded on further, Arf6 and Arl4 GTPases recruit Cytohesin family GEFs to the plasma membrane for increased Arf6 activation implicated in cell proliferation, actin rearrangement, and endocytosis among other processes [31, 32]. In this way, signals from even small amounts of Arf6 activation such as from the Arf6 specific GEF EFA6 would be amplified through downstream effects of the activated GTPase [29, 33].

Finally, current evidence suggests that many GTPases function in crosstalk between pathways often defined by nucleotide-state specific binding. Besides acting in a feedback loop in the GTP-bound state as mentioned above, Arf6 in its GDP bound state recruits the Rac GEF Kalirin5 to the plasma membrane, with subsequent activation of both GTPases leading to cytoskeletal rearrangement and membrane ruffling [34, 35]. And more recently, Rab35 has been shown to promote Arf6 deactivation by recruiting the ArfGAP ACAP1 [36, 37].

Arf GTPases

Arf GTPases are members of the Ras superfamily of small GTPases and include five Arfs and over 20 Arf-like (Arl) and Sar proteins in mammals [10, 38]. Within this larger family, the Arfs fall into three classes, Class I (Arf1/3), Class II (Arf4/5), and Class III (Arf6) based on sequence homology. Arl proteins and most Sar family members are less well characterized and have not been grouped according to sequence or functional homology.

Arfs function in a variety of cellular pathways and are localized throughout the cell (Figure I.4), but share unified functional themes of recruiting coat proteins and activating lipid remodeling enzymes [39-41]. Arf1 functions at Golgi and endosome membranes to initiate vesicle budding by recruitment of coat proteins COPI and Adaptor-Protein (AP-1, AP-3, and AP-4) and GGA protein coat complexes [42, 43]. It also can activate phosphatidylinositol 4- and 5-phosphate kinases (PI4K, PI5K) at the Golgi to aid in actin rearrangement and polymerization [44]. While Arf3, Arf4, and Arf5 all contribute to inter-Golgi transport [41], Arf4 also has a unique role in the development of primary cilia by undergoing an essential interaction with rhodopsin and another protein complex for efficient transport of rhodopsin to the rod outer segment [45, 46].

Arf6 is unique in that it acts as a key regulator in large scale dynamic changes at the plasma membrane such as actin rearrangement, cell migration, cytokinesis, and tumorogenesis [36, 47-50]. One of the most well studied roles of Arf6 is in rapid and widespread endocytosis at the plasma membrane. By enhancing activity of two key lipid remodeling enzymes phospholipase D (PLD) and PI5K, Arf6-GTP enhances both clathrin

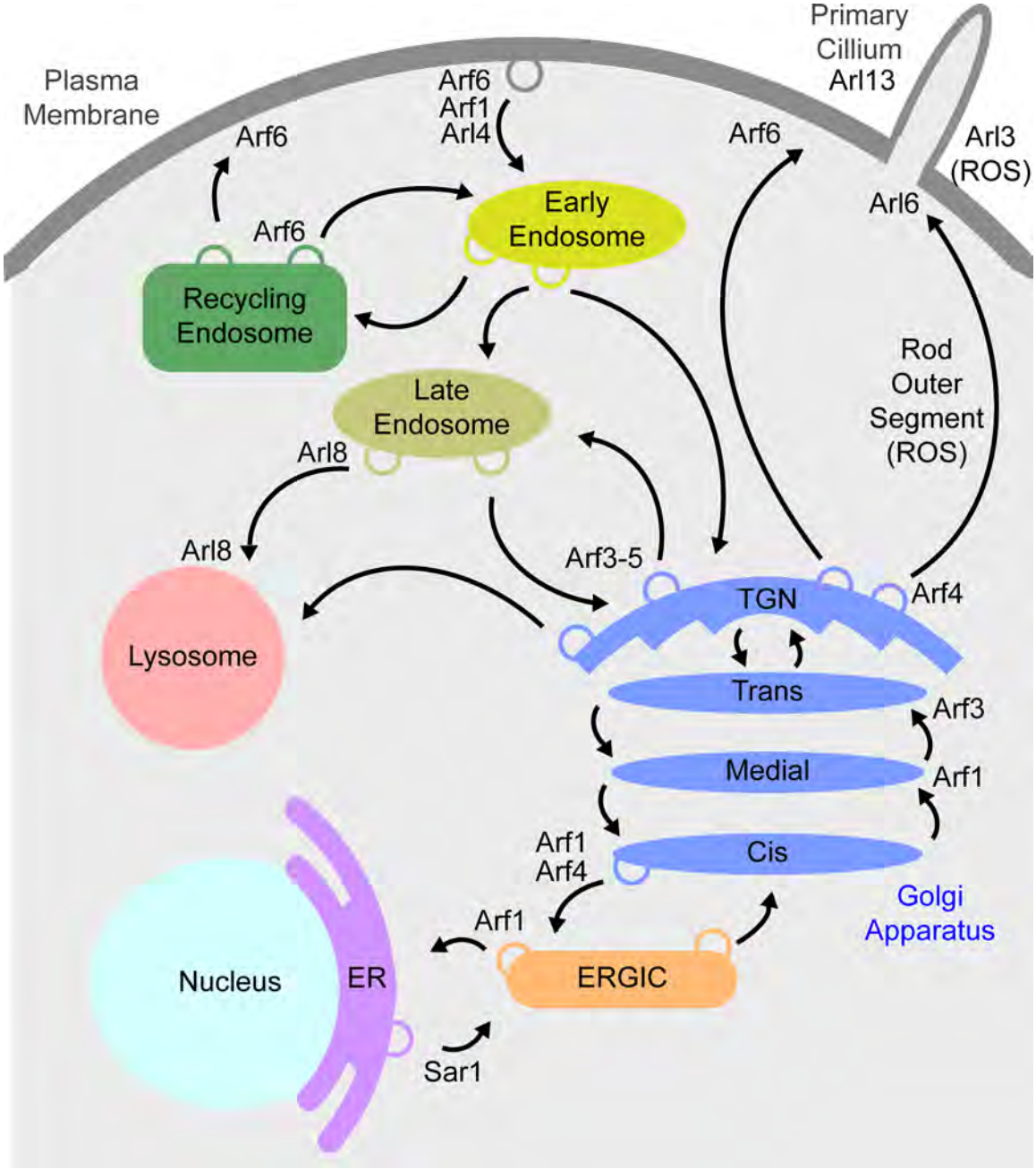

Figure I.4: Cellular Localization of Arf family GTPases. Specific roles of Arf4 and Arl3 in retina cell trafficking to the rod outer segment (ROS) are indicated. ER, Endoplasmic Reticulum; ERGIC, ER-Golgi Intermediate Compartment.

mediated (CME) and clathrin independent endocytosis (CIE) [51-54]. Additionally, Arf6 further regulates CIE through interactions with adaptor proteins, which require deactivation of Arf6 for full internalization of the vesicle cargo. Arf proteins communicate such diverse changes through a conserved mechanism of nucleotide switching.

Arf structural elements and nucleotide binding:

Arf and Arl GTPases consist of a core domain architecture conserved across the Ras superfamily of over 150 proteins (Figure I.5) [55-57]. Within its compact structure is contained a minimal G domain for binding nucleotide and structural 'switch' regions termed switch 1 (SW1), interswitch (ISW), and switch 2 (SW2) flanked by N- and C-terminal regions primarily used to localize the molecule to membrane environments [58]. The G domain includes five consensus motifs spanning the primary sequence that wrap around the guanine nucleotide (Figure I.6A). At the base, the N/TKXD motif coordinates the guanadynyl ring and an alanine in the C/SAK/L/T motif affords specificity for guanine over adenine through main chain hydrogen bonding with the carbonyl oxygen allowed by the small size of the alanine side chain. The 'P-loop' GXXXXGKS/T motif anchors the phosphate groups through hydrogen bonding of the glycine and coordinates a Mg^{2+} cofactor at the S/T residues. The aspartic acid side chain in the DXXG motif coordinates the Mg^{2+} through a water mediated hydrogen bond, and a T consensus motif hydrogen bonds with both the Mg^{2+} and the gamma phosphate of GTP in the active state.

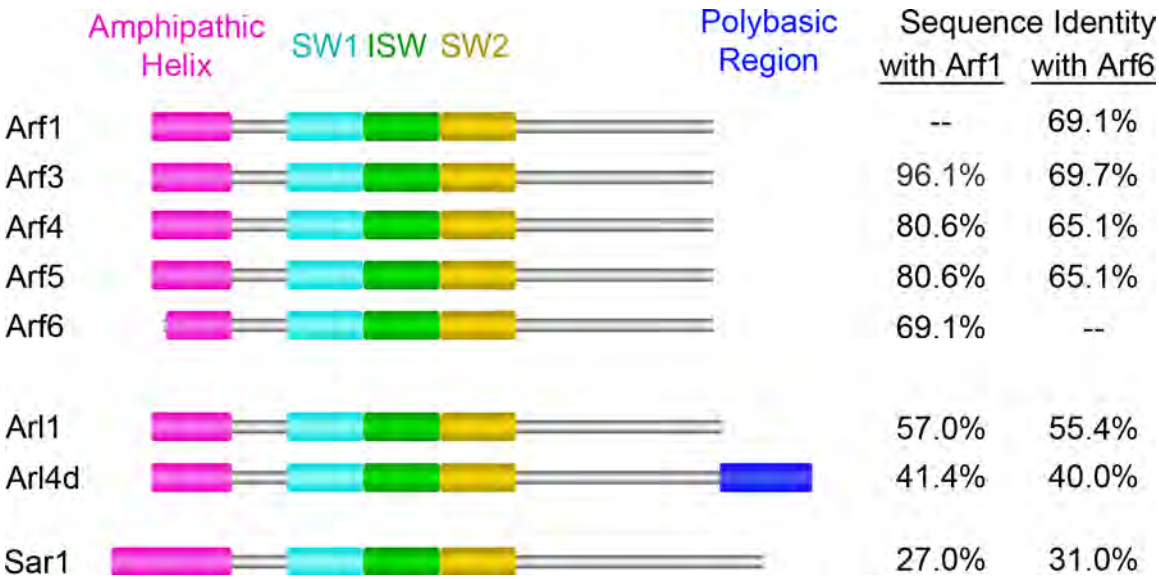

Figure I.5: Domain Architecture and Sequence Conservation of Arf Family GTPases. Switch regions and terminal regions contributing to membrane insertion are colored and labeled as indicated. SW1, Switch I, ISW, interswitch, SW2, Switch II. Sequence identity of family members with Arf1 or Arf6 was calculated with Clustal Omega [56-57].

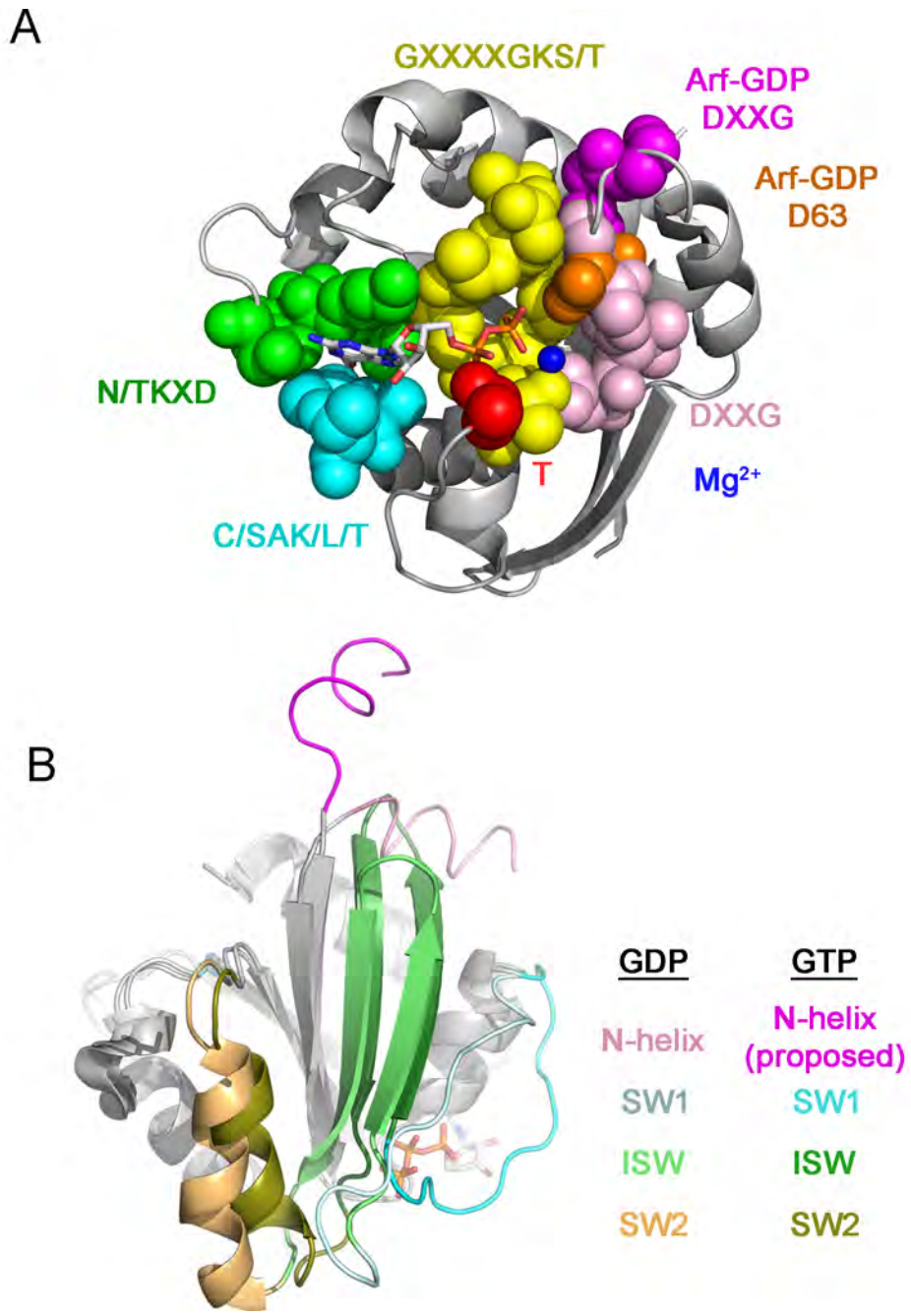

Figure I.6: Arf GTPase Structural Elements A) The nucleotide binding pocket of Arf GTPases. The superimposed Arf6-GTP (PDB ID 2W83) and Arf-GDP (PDB ID 1E0S) structures showing sequence motifs from Arf6-GTP and Mg^{2+}, nucleotide, and DXXG motifs from Arf-GDP. B) Superposition as in A highlighting switch and N-terminal regions.

Included in and adjacent to the nucleotide binding site, the switch regions are so named because they switch between two distinct conformations in GDP and GTP bound forms (Figure I.6B) to propagate the GDP to GTP signal from the nucleotide binding pocket to the rest of the molecule and are thus the sites for effector recognition. They consist of the switch 1 (SW1), interswitch (ISW), and switch 2 (SW2) regions. The Arf family has a largely conserved sequence and structure within the switch regions, with a notable exception being the Arl4/6/7 subgroup containing a five residue insertion in the interswitch region [59].

Structural studies of GDP and GTP bound Arf proteins have elucidated how changes in nucleotide state cause widespread effects through conformational rearrangement [24]. In the Arf6-GDP bound structure, the aspartic acid in the DXXG motif mimics the gamma phosphate of GTP and docks the SWI and DXXG regions in positions incompatible with nucleotide binding. The interswitch and SWII regions are locked into their orientation by a tryptophan clamping interaction with the core. The N-terminal helix docks in a hydrophobic pocket formed at the top interswitch, sequestering it from membrane insertion. Upon binding to GTP, the aspartic acid is reoriented, allowing nucleotide binding of SWI and the DXXG motifs. The SWII region is docked into a new position by a hydrogen bond network coordinated by the tryptophan clamp and an arginine. Interestingly, the ISW undergoes a 3 angstrom (Å) 'interswitch toggle' which promotes an extended conformation competent for membrane insertion. [59].

The change in ISW structure upon nucleotide cycling is unique to Arf family GTPases, as other GTPases show no change in this region as compared in GDP and GTP

bound structures. The interswitch toggle is presumably conserved in all Arf family members except Arl4, Arl6, and Arl7, which have longer ISW regions. Additionally, recent evidence supports the idea that structural rearrangements such as the toggle involved in nucleotide cycling are aided by both effector-mediated as well as spontaneous membrane docking, seamlessly connecting GDP-GTP structural rearrangements and membrane recruitment [60, 61].

Arf Effector Binding

As shown in Figure 1.2, the GTPase cycle determines binding of Arf GTPases with various effectors. These protein interactions by definition do not change the nucleotide state of the Arf, but instead serve to either sequester the GTPase or aid in general signaling pathways through domains outside of the Arf binding motif. Although specific effectors for the GDP bound form of Arf GTPases have been demonstrated [34, 35], the most well studied class is the classical GTP bound effectors. A great deal of structural evidence points to a common hydrophobic area (CHA) on Arfs similar to that observed in Rabs that allows binding with very diverse molecules on the largely conserved surface while maintaining sensitivity to nucleotide states (Figure 1.7a and b) [62, 63]. The CHA (Figure 1.7c) includes key residues at the junction of SW1, ISW, and SW2 which form a deep hydrophobic pocket studded with three conserved residues termed the ‘hydrophobic triad’. Most effector molecules studied to date form lock and key interactions by burying non-polar residue side chains in the pocket. In contrast,

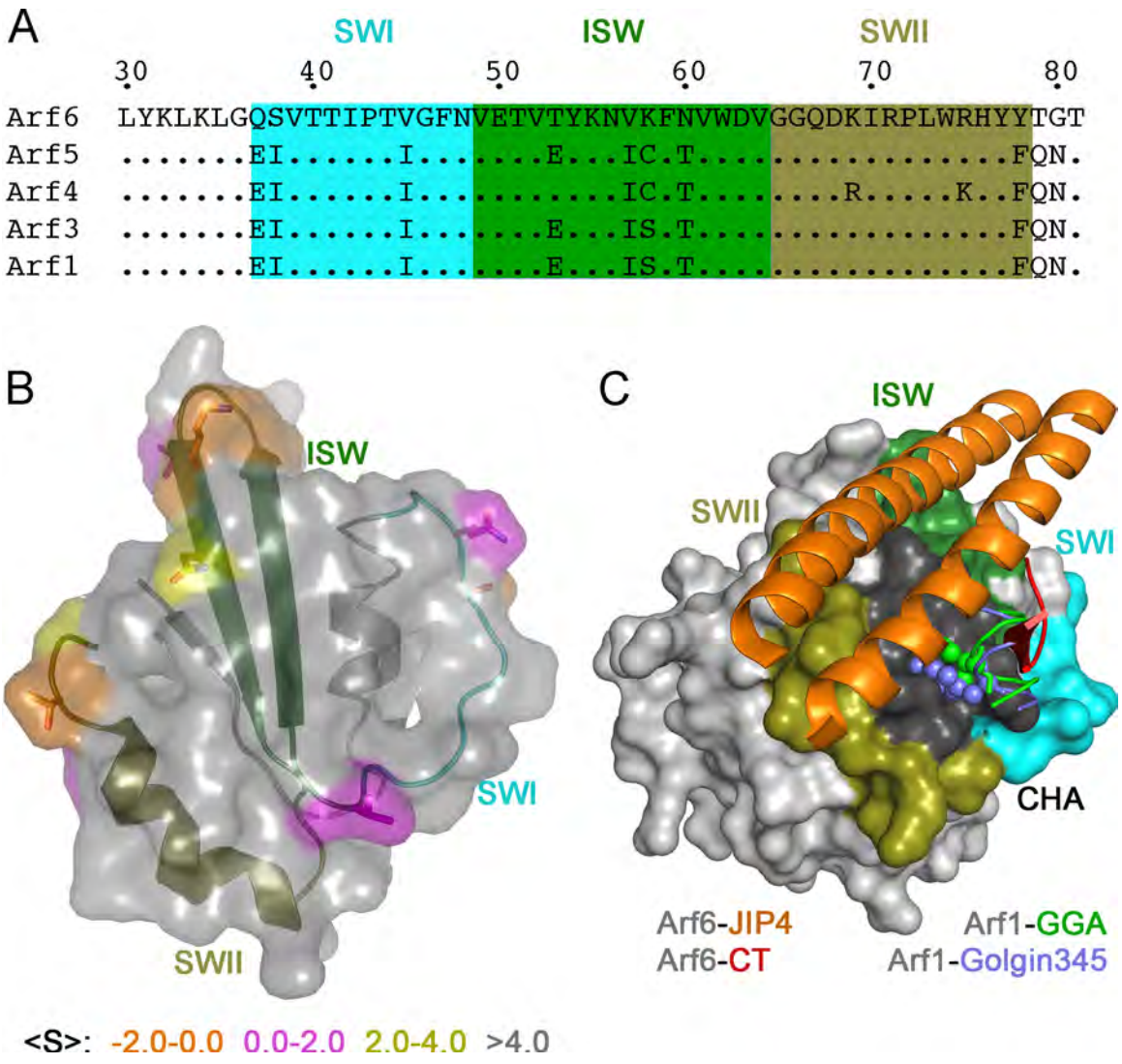

Figure I.7: Arf:Effector Specificity Determinants A) Sequence alignment of Arf isoforms. B) Conservation analysis within Arf switch regions. Structure is Arf6-GTP (from PDB ID 4KAX) with cartoon colored as in A. Surface coloring represents scores in a BLOSSUM62 substitution matrix between Arf6 and Arf1 sequences. C) Structures of Arf-GTP bound to the effectors JIP4, Cholera Toxin (CT), GGA, or Golgin345 proteins. Only relevant portions of effectors are shown for clarity. Arf6 and Arf1 effectors are represented as ribbon/sphere and cartoon, respectively.

epitope of the effector main chain backbone with the triad are not conserved in Arf:effector complexes.

These conserved features in the CHA complement other non-conserved elements shown to be instrumental in selective binding of effectors to particular Arf members. As shown in a conservation map between the two most divergent Arf family members Arf1 and Arf6 (Figure I.7A), of the ten substitutions between these proteins in and proximal to the switch regions, seven residues at the edges of the switch region and in the upper portion of the interswitch β-sheets are plausibly involved in Arf selectivity when considering the key CHA. Specificity determinants have been well documented in the crystal structure of Arf6-GTP in complex with an effector JIP4, important in essential endosome trafficking during the completion of cytokinesis [64, 65]. The dimeric leucine zipper motif of JIP4, lays across the Arf6 switch regions, engaging poorly conserved residues Thr 53, Lys58, Asn60, and Thr79 while excluding interaction with the hydrophobic pocket. Substitution of the analogous residues in Arf1 to the Arf6 equivalents resulted in JIP4 binding equal to that observed with Arf6.

Although these and other results suggest Arf specificity can arise solely from sequence differences [66], it is not clear whether exclusion from the Arf hydrophobic pocket and engagement of non-conserved residues represents a common mode for all Arf6 effectors or a specialized case of Arf6 discrimination. For example, the interaction with Cholera Toxin (CT) from which the Arf proteins were first identified [67] occurs with multiple Arf proteins, and although the Arf6-CT structure identified a binding epitope excluding the hydrophobic pocket similar to the Arf6-JIP4 structure, it is shifted

dramatically away from the nonconserved switch region residues [68]. In general, specificity of Arf effectors has not been studied comprehensively, and some have been shown to interact with specific isoforms through coincidence detection of other lipid or protein interactions [69-74]. As will be expanded upon, specificity within Arf dependent cell signaling events seems to involve a mixture of strict biochemical selectivity for specific protein-protein recognition as well as colocalization of the proteins at membranes.

Membrane Recruitment of Arf GTPases

Arf family GTPases associate with membranes through regions flanking the core G domain (Figure I.5). The N-terminus consists of an amphipathic helix that can orient into membrane bilayers based on hydrophobicity. Additionally, all Arfs and many Arls are myristoylated at the 2-Gly position, and other Arl proteins are acetylated [10]. In Arfs, this modification aids in docking of the helix into the bilayer through insertion of hydrophobic sides chains and myristoyl moiety [23]. This association is dynamic within the membrane space, and allows for independent movement of the N-terminus and G domain [60, 61]. In addition, some Arl proteins also contain C-terminal basic regions for membrane attachment. All isoforms of Arl4, for example, localizes to the plasma membrane in the active state, and mutation of the myristoylation site or truncation of the C-terminal polybasic region decreased this effect [75].

Arf membrane recruitment is coupled to nucleotide state, and in all cases studied to date Arf-GTP is more stably associated with membrane environments than Arf-GDP.

This fact coupled with the discovery of a buried N-terminus in the Arf–GDP crystal structures originally led to the idea that Arf GTPases were strictly excluded from membranes until activated [23, 76]. However, key biochemical, cell biological, and structural experiments have shown that the amphipathic helix can mediate association with membranes in the GDP-bound state [77]. Recently, nuclear magnetic resonance (NMR) studies of Arf1 bound to lipid bicelles have clarified association determinants [60, 61]. In the GDP-bound structure, the N-terminal amphipathic helix adopted a highly flexible structure and showed measureable bilayer association even in the absence of myristoylation. Increased association occurred in the GTP-bound state coinciding with a more rigid position of the slightly elongated helix, although approximately 90° mobility was still possible**.** Importantly, membrane binding of Arf-GDP brought about structural changes in the switch regions. Additionally, a groove speculated to be the myristoyl group binding pocket in the Arf-GDP structure was covered in the GTP-bound structure by residues in the interswitch loop undergoing the toggle register shift during nucleotide exchange. Together this evidence suggests that nucleotide exchange and membrane recruitment follow concerted rather than sequential mechanisms and define the role for amphipathic helix insertion in membrane recruitment.

Arf GEFs and GAPs

GTPases undergo intrinsic GDP-GTP exchange and hydrolysis that is accelerated by GEFs and GAPs, respectively. Arf family GTPases are activated by the Sec7 family of GEFs, which all have a conserved mechanism for nucleotide release [78, 79]. While

there are 15 Sec7 family GEFs, there are over 30 mammalian Arf GAPs. Arf GEFs and GAPs can be broadly grouped into large and small members based on size and architecture. The small GEFs have a simplified architecture with minimal features outside of the catalytic domain, and large family members typically include diverse regulatory features essential for coordinating and regulating cell signaling at sites of action.

Sec7 Domain GEFs:

Sec7 domain GEFs are so named due to the presence of a catalytic domain analogous to the Sec7p in yeast [80, 81]. The Sec7 family consists of GBF/BIG, EFA6, BRAG/IQ, Fbox, and Cytohesin subfamilies [78]. Although all of these proteins studied to date have measured GEF activity, the specificity of these enzymes for the Arf isoforms is not well understood. EFA6, for example, has clear preference for Arf6 over Arf1, whereas Cytohesin GEFs may be quite promiscuous *in vivo*, despite showing higher activity for Arf1 in solution [33, 82]. These GEFs may activate specific Arf isoforms through colocalization at membrane compartments via regulatory protein or membrane binding domains.

Sequence conservation and structural studies of Sec7 domains have shown a highly conserved α-helical protein fold [59, 83-88]. The Arf interface can be separated into a seven member superhelix binding the SW1 region and a three helix bundle binding the ISW and SW2 regions (Figure I.8). Structural studies using the Cytohesin family member ARNO in intermediate stages of nucleotide exchange have revealed the mode of

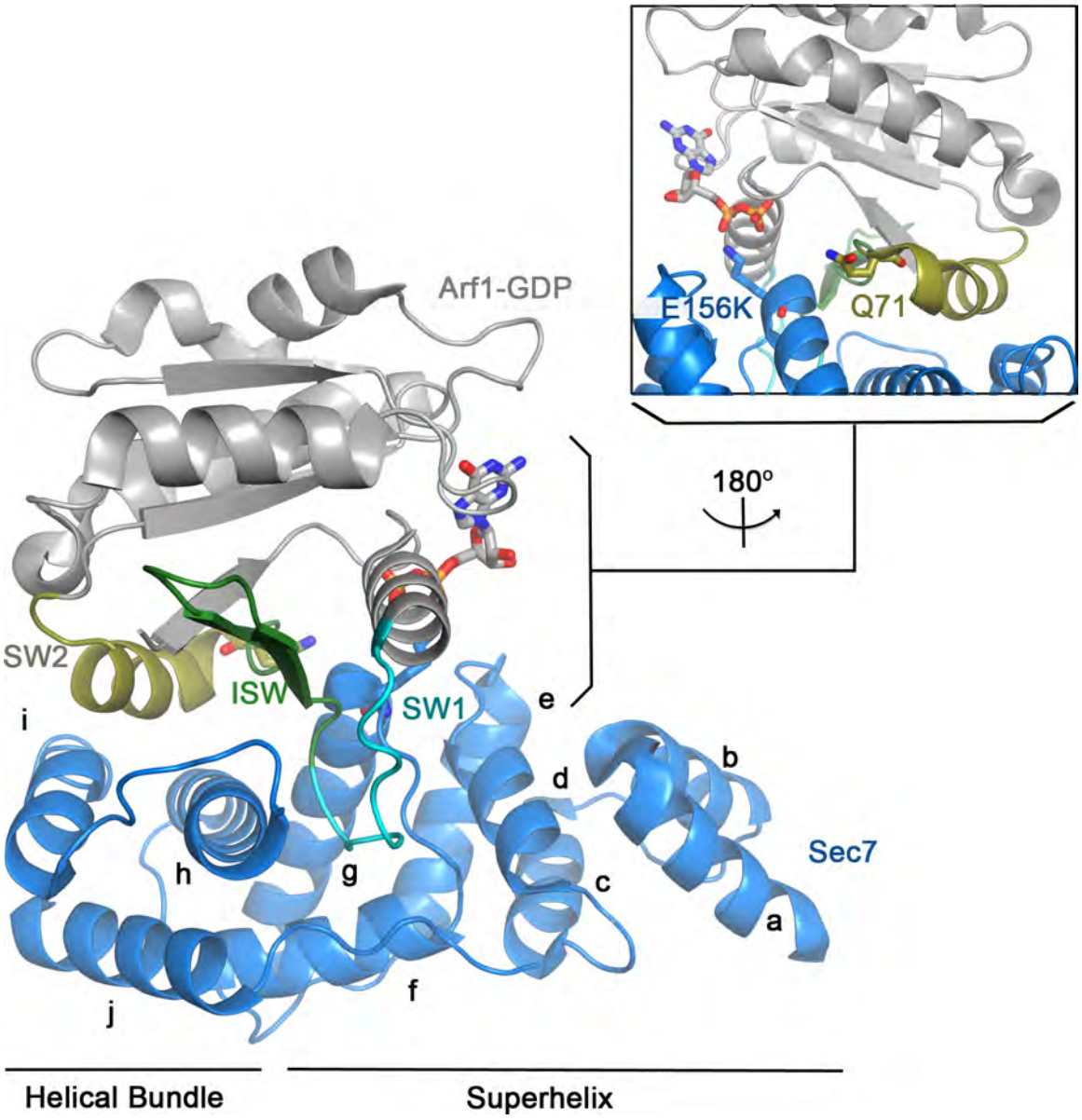

Figure I.8: The ARNO Sec7 Domain in Complex with Arf-GDP. The structure of the ARNO Sec7 domain in complex with Arf1-GDP (PDB ID 1R8S). Regions of interest are highlighted as indicated. The mutation of the catalytic Glu in the Sec7 domain and the catalytic Gln for hydrolysis in Arf1 are represented in sticks.

interaction with substrate Arf and the catalytic mechanism by which GDP is released [79, 89]. Arf-GDP forms an extensive interface with the Sec7 domain in which the interswitch region adopts its GTP-bound conformation. A key structural element termed the glutamic acid finger inserts into the β-phosphate site on the Arf molecule. Mutation of this residue to a lysine abolishes GEF activity by stabilizing the Arf-GDP complex [89]. Additionally, a Sec7 αF/αG loop aids in nucleotide release by competitive binding to SW1 residues involved in phosphate and Mg^{2+} binding. Displacement of the nucleotide is stabilized by large rearrangements in the switch regions favoring a nucleotide free state of the Arf [79]. After release of the GDP, the excess of GTP in the cytosol favors completion of GDP-GTP exchange.

GAP catalysis

Arf GAPs fall into ten classes all containing a canonical GAP domain [90, 91]. Although structural evidence for these proteins is limited, structures of the catalytic domains alone or in complex with Arf-GTP or transition state mimetics have shown a conserved domain structure but potentially diverse catalytic mechanisms [92-95]. An initial structure of ArfGAP1 in complex with the Arf1-GDP showed the catalytic domain only contacting the SW2 region of the substrate, and no catalytic residues in the nucleotide binding pocket, suggesting a mechanism different than Ras GAPs whereby the GAP accelerates hydrolysis by stabilization of SW2 possibly in association with other proteins [95]. In contrast, the structure of ASAP3 with Arf6 in the GDP-ALF_3 transition state mimetic showed a flipped orientation of the GTPase with full engagement of both

the SW1 and SW2 regions and an ‘arginine finger’ protruding into the nucleotide binding site. Unlike other GAPs however, an additional aspartic acid acts to stabilize the Arf DXXG motif and facilitate phosphate cleavage, with the arginine stabilizing the transition state. Finally, ASAP3 and other ASAP family members were shown to be regulated by Ca^{2+} interactions at SW2 [92]. This suggests a possible divergence of catalytic mechanisms among Arf GAP members.

Studies to date regarding the substrate recognition by the ArfGAP domain have been complicated by exclusion of external domains known to play key roles in functional specificity *in vivo*. Examples include ankirin repeats in the ACAP subfamily of Arf GAPs that bind vesicle cargo, COPI binding sites in ArfGAP1, and PH and BAR domains that bind membranes in the ASAP subfamily [90]. It is therefore not surprising that, as with Arf GEFs, only weak determinants of Arf-GTP substrate specificity in Arf GAPs have been demonstrated. As is the case with GEFs, functional specificity of Arf GAPs *in vivo* may often be controlled as much by colocalization via protein-protein or protein-lipid interactions of domains outside the catalytic domain as by minor sequence differences in the catalytic core.

Membrane Binding Domains in GEFs and GAPs

Whereas localization to membranes for GTPases is primarily through external lipid attachment and membrane insertion, GEFs and GAPs often couple membrane insertion elements like basic helices with large membrane binding domains (MBDs). A consensus is developing that binding of MBDs and other insertion elements directly

contribute to the organization of trafficking events not only by recognizing low abundance membrane components that are upregulated at specific locations or times but also by sequestering lipids and defining the polarity of vesicles. Additionally, multi-domain GEFs and GAPs often serve as platforms for other trafficking molecules, enhancing the natural accumulation due exclusively to membrane components. Moreover, MBDs act in the physical process of membrane deformation by displacing outer leaflet lipids to establish curved membrane vesicle precursors [96, 97]. Phosphatidyl inositol (PtdIns) (4,5)-diphosphate binding of ENTH domains has been directly linked to membrane tubulation, defining a possible role for these proteins in endocytosis at the plasma membrane [98]. BAR domains, such as the one found in the ArfGAP ASAP1, and the ALPS motif in ArfGAP1 use nonspecific curvature-dependent interactions for membrane recruitment [99-101].

MBDs involved in membrane trafficking have varied membrane recruitment mechanisms. Some domains such as C2, use nonspecific membrane insertion of basic regions into the polar surface of the membrane or hydrophobic regions into the hydrocarbon core. In these types of proteins, membrane insertion can be regulated by calcium flux. C1 domains specifically recognize diacylglycerol (DAG) and phorbol esters, and are thought to be stabilized by interactions with acidic lipid head groups. A large majority of membrane binding domains like PH, FYVE, PX, and ENTH domains can include recognition sites for low abundance phosphoinositides [102-105]. Each domain's structure is evolved for specific function in membrane trafficking. Since the following work focuses on the cytohesin family of Arf GEFs which contain a PH domain,

the proceeding will detail this particular domain's phosphoinositide recognition and provide relevant context for further discussion.

Phosphoinositide Recognition in PH domains

The Pleckstrin Homology (PH) domain is a large family of MBDs, with over 250 proteins involved in cellular signaling throughout the cell. They consist of a β-barrel structure of seven β-strands with an α-helical 'cap' at one end (Figure I.9). Although this basic fold is highly conserved, there is little sequence identity between family members. Phosphoinositide binding is mediated through a large electropositive groove in the open end of the β-barrel. A conserved $KX_n(K/R)XR$ motif including residues in and adjacent to the β1/β2 loop interacts with one or more phosphate groups. With such a general mode of recognition, and little sequence conservation, it is not surprising that the majority of PH domains studied show low selectivity and bind any one of the seven occurring possible mono-, bis-, or tris-phosphorylated phosphoinositide species with equal affinity [106, 107]. Most PH domains, and MBDs in general, supplement weak binding with structural determinants outside the phosphoinositide binding pocket. In some cases however, robust membrane targeting can be achieved almost completely through the increased phosphoinositide recognition or direct membrane insertion of the PH domain, owing to hypervariable loops extending away from the conserved barrel, β1-β2, β3-β4, and β6-β7 loops [106]. Structural studies of particular highly discriminating PH domains have been critical to elucidate the mechanisms by which they recognize polyphosphoinositides such as PtdIns $(3,4)P_2$, PtdIns$(4,5)P_2$, or PtdIns$(3,4,5)P_3$.

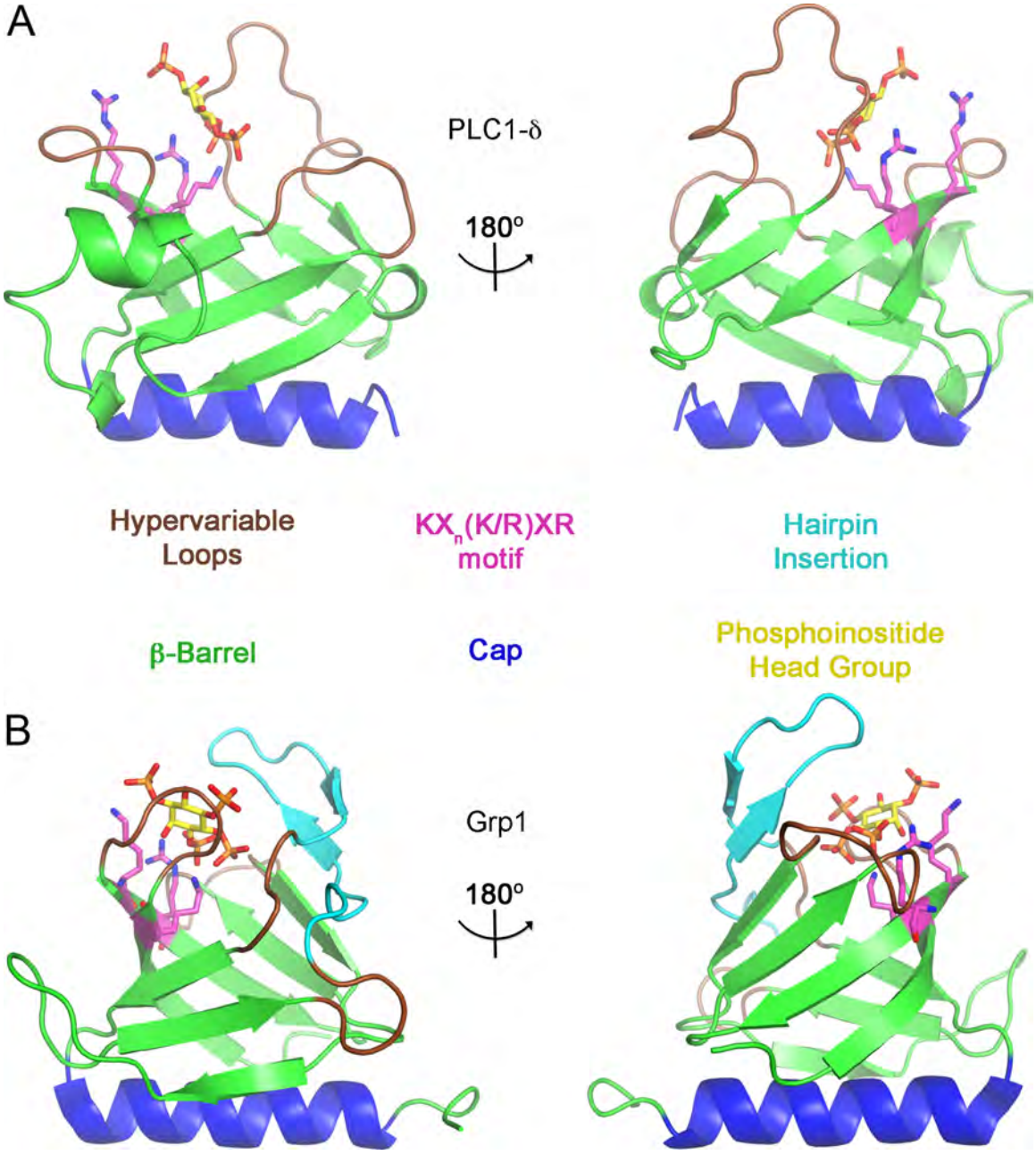

Figure I.9: Overall Structure of Representative PH Domains. Overall structures of PLC1-δ (A) and Grp1 (B) depicting structural elements as indicated.

A wide variety of phosphoinositide recognition modes have been observed in these PH domains (Figure I.10). The PLC1-δ PH domain achieves high affinity binding through interactions between the 4- and 5-phosphates and Lys 32, Lys 57, and Asn 106 from each of the three variable loops, forming a structure that precludes 3-substituted phosphoinositides (Figure I.10A) [108]. TAPP1 seems to bind PtdIns(3,4)P_2 through similar means, but lacks any contribution from its β6-β7 loop and has a key alanine side chain to block 5-substituted species [109] demonstrating specificity through selectivity rather than affinity. The related protein DAPP1 conversely allows binding of PtdIns(3,4,5)P_3 through a substitution of this alanine to a glycine (Figure I.10B) [110].

In addition to these studies, the presence of unique phosphoinositide recognition beyond simple affinity differences has been established in structural studies of PH domains from the cytohesin family Arf GEFs Grp1 and ARNO. The phosphoinositide binding profiles have been exhaustively studied, and crystal structures have been obtained for multiple conserved family members bound to different species of phosphoinositide head groups. Binding studies have shown clear preference for PtdIns(3,4,5)P_3, which is negated in an alternately spliced form of these proteins resulting in a single insertion of a glycine at the beginning of the β1-β2 loop [82, 111-113]. These forms are referred to as the diglycine (2G) and triglycine (3G) forms. Crystal structures of 2G and 3G PH domains bound to Ins(1,4,5)P_3 or Ins(1,3,4,5)P_4 helped explain these findings.

High affinity selective binding of PtdIns(3,4,5)P_3 is achieved through a large extension of the β6-β7 loop referred to as the hairpin insertion (βi1-βi2). The crystal

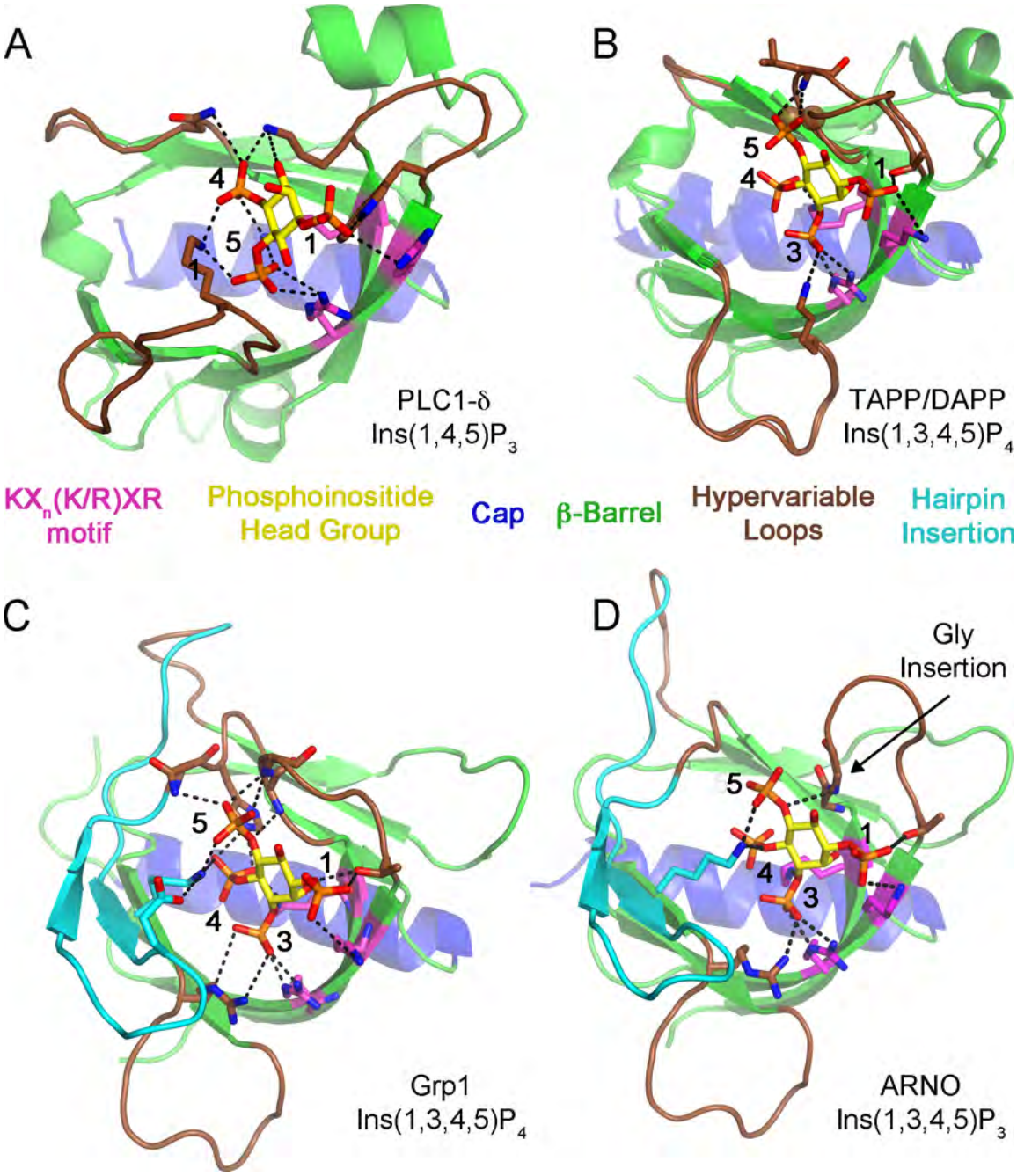

Figure I.10: Phosphoinositide Recognition in PH Domains. A) PLC1d structure bound to Ins(1,4,5)P_3 (PDB ID 1MAI). B) TAPP1 (PDB ID 1EAZ) and DAPP (PDB ID 1FAO) structures superimposed and showing the Ins(1,3,4,5)P_4 bound in the DAPP/TAPP structure. The TAPP1 alanine key for selectivity is shown in spheres. C) Grp1 structure bound to Ins(1,3,4,5)P_4 (PDB ID 1FHX). D) 3G ARNO structure bound to Ins(1,3,4,5)P_3 (PDB ID 1U27).

structures of the 2G Grp1 and 3G ARNO in complex with Ins(1,3,4,5)P_4 show interactions between side chains in the basic motif, β1-β2 loop, and β3 strand with the 1-, 3-, and 4-phosphates. In addition, the presence of the hairpin insertion along with key threonine and arginine residues from β1-β2 and β3-β4, respectively, helps create an extensive network of interactions to cradle all three substituted phosphates in the binding pocket. Of particular note are interactions of Lys 343 and His 355 in β6-β7 with the 5-phosphate. In the 3G form, the extra glycine in the β1-β2 loop causes an overall structural change resulting in a loss of main chain hydrogen bonds with the 5-phosphate of PtdIns(3,4,5)P_3 (Figure I.10C and D) [111, 114, 115].

Despite the fact that structural differences between 2G and 3G PH domains affect 5-phosphate interactions seen in many other PH domains bound to di-substituted phosphoinositides, the net effect is only to reduce affinity for PtdIns(3,4,5)P_3, as Ins(1,4,5)P_3 and Ins(1,3,4)P_3 are recognized equally well in *in vitro* experiments [111]. This led to exploration of other possible binding geometries in phosphoinositide binding. A crystal structure with a 3G PH domain in complex with Ins(1,4,5)P_3 showed rotated orientations of the phosphoinositide head group when compared to Ins(1,3,4,5)P_4 to accommodate increased interactions away from the β1-β2 loop (Figure I.11A). This organization of the binding pocket would also allow Ins(1,3,4)P_3 to dock in a 'flipped, rotated' position [111]. Using the symmetric Ins(1,3,4,5)P_4 orientation seen in multiple PH domains as a starting model, one can recognize that PLC1-δ adopts a 'flipped' position, suggesting that PH domains have structurally diverged for specialized phosphoinositide binding properties but maintain generalized functions (Figure I.11B).

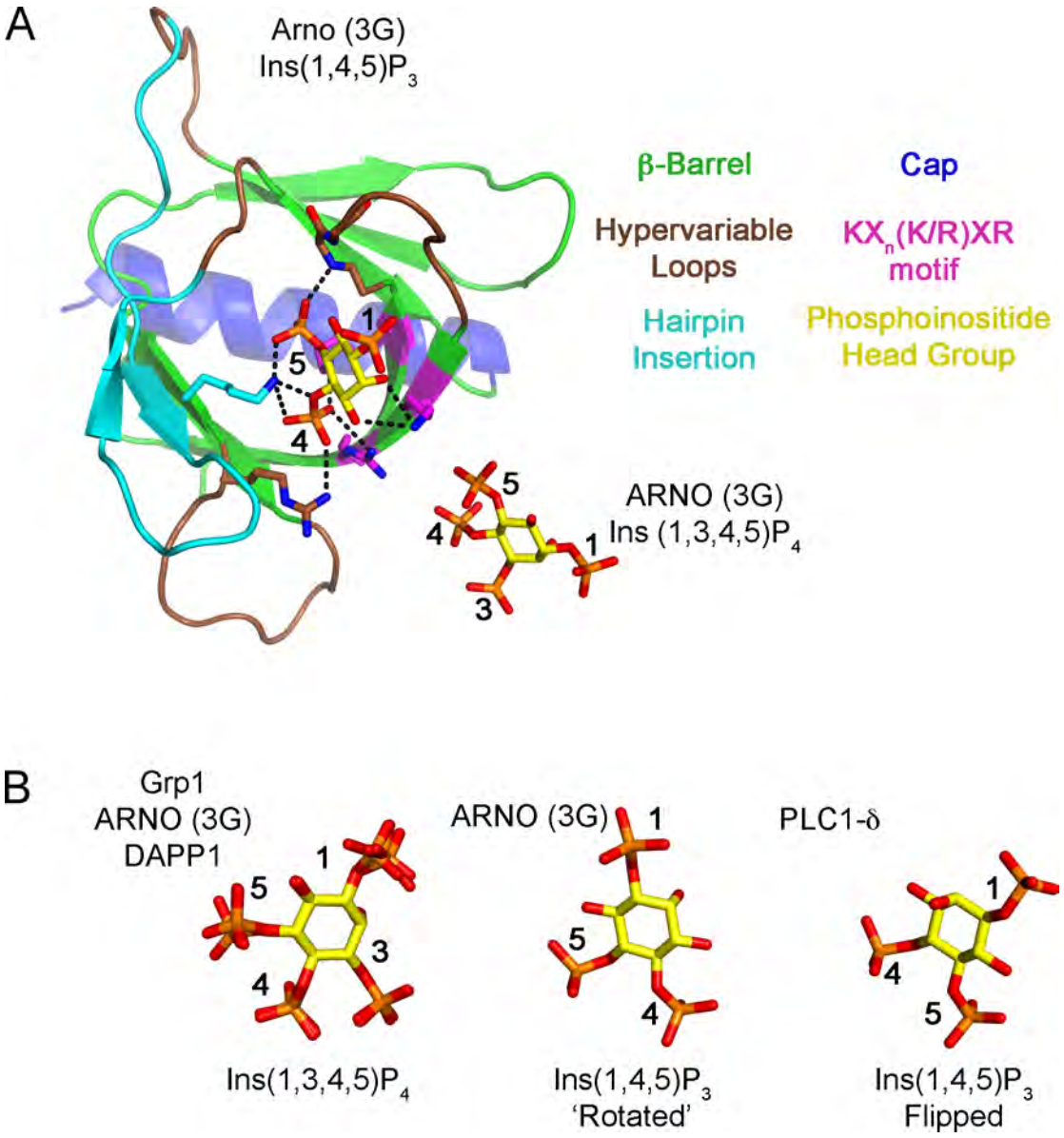

Figure I.11: Variable Phosphoinositide Binding Geometries in PH domains. A) Crystal structure of the 3G ARNO PH domain in complex with Ins(1,4,5)P_3 (PDB ID 1U29). For comparison, the Ins(1,3,4,5)P_4 head group from the 3G ARNO PH domain structure (PDB ID 1U27) is shown offset after superposition of the two structures. B, left) Ins(1,3,4,5)P_4 in Grp1 (PDB ID 1FHX), DAPP1 (PDB ID 1FAO), and ARNO (PDB ID 1U27). B, middle) (Ins1,4,5)P_3 in ARNO. B, right) PLC1-d (PDB ID 1MAI). Structures are shown after superposition of PH domains and offset for comparison. Grp1, DAPP1, and ARNO Ins(1,3,4,5)P_4 molecules are overlaid.

Membrane Insertion by PH domains

Although many PH domains can localize to membrane environments through phosphoinositide head group binding, phosphoinositide binding is weak or indiscriminate in most cases [107], such that phosphoinositide binding alone is insufficient to recruit these PH domains to membranes. Indeed, even in the case of high affinity binders, the low abundance of phosphoinositides even in the presence of stimulating factors like insulin necessitates additional factors for association. In at least some cases PH domains are localized through protein-protein interactions as well. For example, FAPP1 PH domain simultaneously binds PI(4)P and Arf-GTP [73, 74]. A less well recognized mode of PH domain membrane recruitment is direct membrane insertion independent of but complementary to phosphoinositide recognition. Two examples well-studied in computational and biophysical experiments are the proteins Akt and Grp1, both of which have high affinity for PtdIns(3,4,5)P_3.

Phosphoinositide binding in the Akt PH domain is dependent on a 'sentry' glutamate located in the β1-β2 loop that forms a bridge across the phosphoinositide binding pocket [116, 117]. A Glu to Lys mutation of this residue has been linked to ovarian cancer [118] by increasing PtdIns(4,5)P_2 affinity, leading to rampant activation of the Akt pathway involved in cell proliferation. Structural rearrangement of the glutamate is thought to aid in its direct insertion into the membrane. A sentry glutamate was also identified in Grp1, where membrane insertion seems to be aided by the extra hairpin insertion in the β6-β7 loop [119]. Fluorescence Resonance Energy Transfer (FRET)

experiments showed association to anionic membranes containing phosphatidylserine but absent of PtdIns(3,4,5)P_3 later confirmed by microscopy and biophysical measurements, pointing to interactions of the β1-β2, β3-β4, and β6-β7 barrel loops in membrane docking [120-123]. This was supported in molecular dynamics simulations and detected directly through monolayer penetration, dependent on portions of the all three Grp1 variable loops and enhanced by PtdIns(3,4,5)P_3 [124]. These findings suggest a mechanism whereby high PtdIns affinity PH domains use nonspecific insertion of their loops in a search for rare lipid bilayer species. Insertion additionally anchors domains beyond recognition of single lipid species [125]. Although many PH domains do not have large extensions like Grp1, membrane insertion may still play a role in stabilizing membrane association after identification of a target phosphoinositide.

Cytohesin GEFs

Family Expression, Function, and Domain Organization

The small Arf GEFs of the cytohesin family consists of Cytohesin-1 (CYTH1), ADP Ribosylation Factor Nucleotide Site Opener (ARNO), General Receptor of Phosphoinositides (Grp1), and Cytohesin-4. Of these only CYTH1, ARNO, and Grp1 are ubiquitously expressed, with Cytohesin-4 expressed only in blood leukocytes [126]. Additionally, expression of the diglycine and triglycine splice variants described earlier is family member specific. CYTH1 and ARNO are expressed predominantly in the triglycine form, whereas Grp1 and Cytohesin-4 are expressed in the diglycine form.

Cytohesins are recruited to the plasma membrane upon receptor stimulation due to upregulation of phosphoinositol-3-OH-kinase (PI3K). This finding combined with their role as GEFs for Arf GTPases implicates them as major players in many cellular responses such as cytoskeletal rearrangement, receptor endocytosis, and receptor trafficking at recycling endosomes, all of which are dependent on levels of active Arf6 and PtdIns(3,4,5)P_3. Independent of Arf6 however, connections have been made between Cytohesins and insulin resistance by studying the Drosophila cytohesin *steppke*. This homolog is essential for proper downstream insulin signaling, and knockdown or chemical inhibition of cytohesin function results in growth defects and insulin resistance [127, 128].

The Cytohesin domain architecture consists of an N-terminal Heptad Repeat (HR) coiled-coil domain, a catalytic Sec7 domain, a Pleckstrin Homology (PH) domain with splice variant dependent specificity for the phosphoinositide PtdIns(3,4,5)P_3 and PtdIns(4,5)P_2 or PtdIns(3,4,5)P_3 discussed earlier, and a C-terminal helix (CtH) including an 8-10 residue polybasic region (pbr). Although the Sec7 and PH domains perform essential, well-defined roles in Arf exchange and membrane targeting, the heptad repeat may also contribute as a scaffolding domain. As will be expanded upon later, the pbr helps form a complex regulatory mechanism for these processes from only minimal structural components.

Heptad Repeat Coiled-Coil Dimerization

The N-terminal HR region of cytohesins has been shown to interact with coiled coil modules in other proteins such as Grsp1 (Grp1 Signaling Protein), GRASP1 (Grp1 Associated Scaffolding Protein), and CASP (Cytohesin-1 Associated Scaffolding Protein) [129-132]. These heterodimeric interactions, as the protein names imply, may help localize cytohesins to specific membrane-protein environments or organize higher-order complexes.

The heterodimeric coiled-coil interaction of Grsp1 is the most well-studied of the three and is mediated through a heptad repeat region next to a FERM domain shown in some cases to interact with cytoplasmic tails of transmembrane proteins as well as PtdIns(4,5)P_2 [129]. Grp1-Grsp1 dimers have been immunoprecipitated from cells and localize to the plasma membrane upon insulin stimulation. In *in vitro* experiments, Grsp1 does not form homodimers, but heterodimers are readily formed with CYTH1 and Grp1 in biochemical experiments, whereas ARNO, the most divergent family member, showed limited heterodimer formation [130].

Interactions between Cytohesins and the CASP and GRASP proteins are less well characterized, but have been identified through yeast two hybrid screens and characterized in *in vitro* binding and cell biological assays [131, 132]. Like the Cytohesin-Grsp1 interaction, these interactions occur on the plasma membrane, but they may have different cellular determinants. Cytohesin recruits CASP from Golgi compartments to membrane ruffles at the PM upon Epidermal Growth Factor stimulation.

Conversely, GRASP1 recruits Cytohesin to the plasma membrane independent of stimulation [132].

Interestingly, ARNO and Grp1 also form less well-characterized homodimers to potentially concentrate Cytohesin localization upon phosphoinositide signals, however the relative stability of homo- and heterodimeric complexes varies between Cytohesin family members and scaffolding partners [130]. Observations *in vitro* showed that phosphoinositide-dependent membrane recruitment as well as GEF activity was independent of dimer formation [133]. More recently, the ARNO coiled–coil domain has been shown to interact with its PH domain in cis at an Akt phosphorylation site. Disruption of this interaction through a pseudophosphorylation mutant resulted in increased plasma membrane recruitment [134]. Together, these data suggest that coiled-coil interactions along with specific PH domain recruitment in Cytohesins may work competitively or cooperatively for variable recruitment under variable conditions of hormone stimulation or membrane composition, potentially offering fine-tuned regulation as well as levels of constitutive localization for homeostatic functions. Additional studies are needed to clarify the physiological role of homo- and heterodimeric complexes.

Autoregulatory Interactions in Guanine Nucleotide Exchange Factors

Autoregulation is a common mechanism in which the catalytic activity of an enzyme is controlled through inhibitory intramolecular interactions, typically relieved by post-translational modification or effector binding at an allosteric site. This phenomenon has been seen in transcription factors, kinases, and actin organization proteins among

others [135]. A number of GEFs have shown structural autoregulation [26, 87, 133, 136-140]. Two well-documented but distinct examples in Ras and Rho GTPase signaling are Son of Sevenless (Sos) and Vav1. In the ArfGEF field, autoregulation in the presence or absence of membranes is a common theme, as evidenced by the yeast protein Sec7 and mammalian protein BRAG2. Cytohesin GEFs have an intricate mechanism of autoinhibition using minimal architecture. Together, these GEFs regulate exchange activity and membrane localization signals through diverse mechanisms.

Son of Sevenless and Vav1: Autoregulation in Ras and Rho-GTPases

Sos is a Ras GEF with Histone, Dbl Homology (DH), PH, Rem, and Cdc25 domains with complex autoregulation that incorporates membrane recruitment with catalytic activation. Sos acts in a feedback loop in which Ras-GTP stimulates exchange of Ras-GDP to Ras-GTP. Ras-GTP binds an allosteric site, driving large-scale conformational changes allowing access to the catalytic Cdc25 domain by Ras-GDP [30]. Structurally, the DH and PH domains block the allosteric site on the catalytic Dbl domain [138] and intriguingly, the histone domain contributes to this autoinhibition by stabilizing the DH-PH module [141] (Figure I.12A). Autoinhibition relief is directly coupled to membrane recruitment. As shown from extensive structural and biochemical studies, histone domain binding to acidic membrane lipids and PH domain binding to PtdIns(4,5)P_2 allows rearrangement of all three inhibiting domains and in turn accessibility of both the allosteric and catalytic binding sites on Sos to Ras [141, 142].

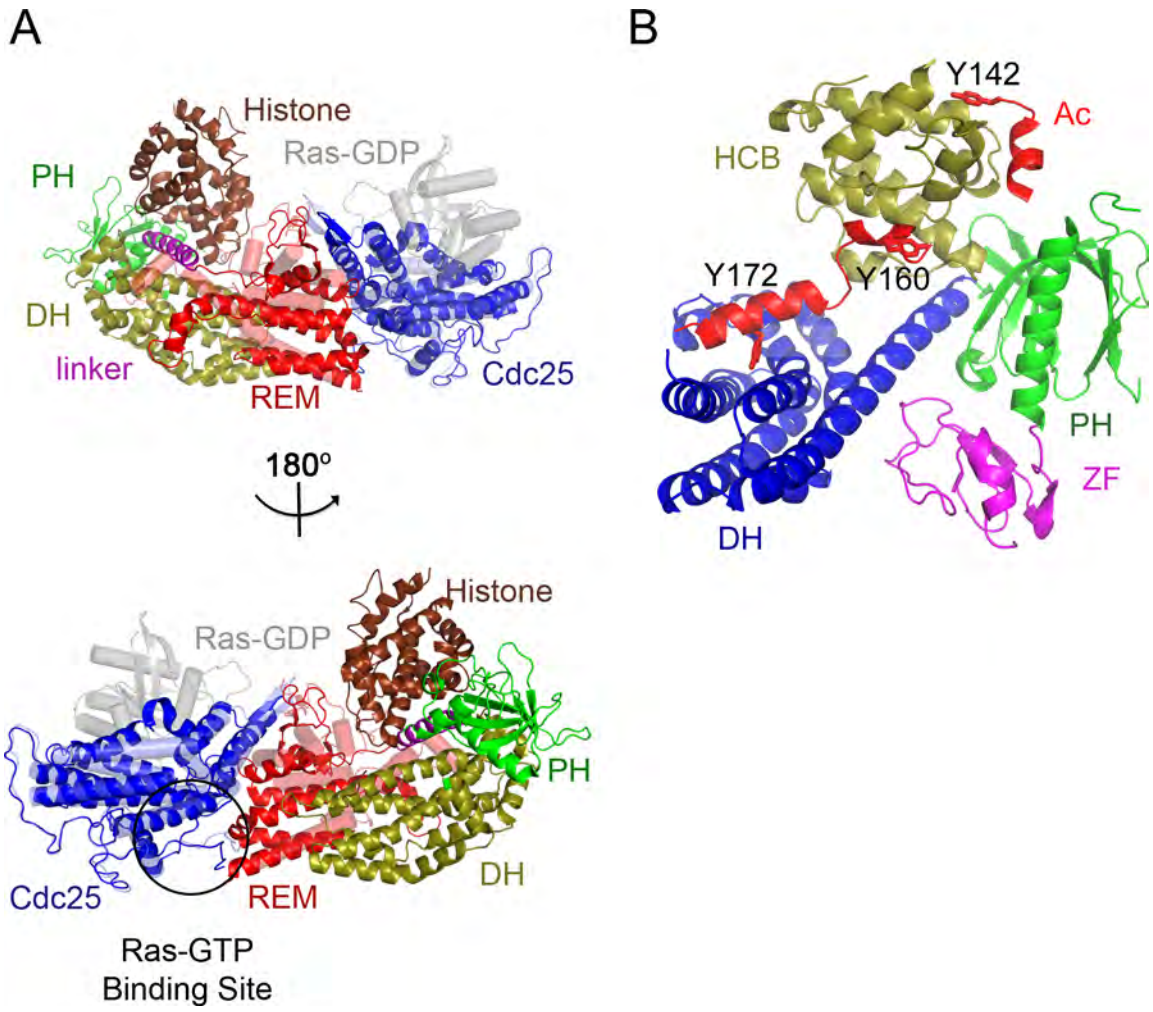

Figure I.12: Autoregulation in Sos and Vav1. A) Structure of Son of Sevenless (Sos) in autoinhibited (PDB ID 3KSY, cartoon helices) and Ras-GTP/Ras bound (PDB ID 1XD2, transparent with cylindrical helices) conformations. The Ras-GTP molecule has been eliminated for clarity and the Ras-GTP location when bound to the REM domain is outlined. Note steric clashes between the Histone/DH/PH tandem in the autoinhibited structure and the active conformation of the REM domain. B) Autoinhibited Vav1 structure (PDB ID 3KY9). Tyrosines that are substrates for Src kinase are shown in sticks and labeled.

Vav1 is an oncogenic RhoGEF with a complex domain architecture resulting in multiple layers of autoregulation. The catalytic DH domain is blocked by an N-terminal acidic domain (Ac) that docks into the Rho-GDP binding site (Figure I.12B) [143, 144]. Phosphorylation of the Ac domain at Tyr 174 by Src kinases releases the helix and facilitates catalysis. This first Ac-DH interaction is stabilized by a calponin homology (CH) domain with the DH, PH, and Ac domains, which promotes a completely closed conformation [145]. Two additional Src phosphorylation steps at Tyr 147 and Tyr 160 release this first interaction. Together this data show stepwise inhibition of GEF activity dependent on phosphorylation but independent of membrane environments, in contrast to Sos.

Autoregulation in the ArfGEFs Sec7 and BRAG2

The yeast ArfGEF Sec7 is the founding Sec7 domain family member first identified in Golgi traffic regulation and the sole ortholog of the BIG1/2 family of mammalian Arf GEFs [80, 146]. Autoinhibition stems from interaction of the GEF domain with the C-terminal HDS1 domain and is relieved by Arf1-GTP interaction with the HDS1 domain, which is sufficient for recruitment to Golgi membrane environments independent of the presence of Golgi phosphoinositide PtdIns(4)P (Figure I.13A). Other portions of the protein, such as the N-terminal DCB and HUS domains or the C-terminal HDS 2-4 may relieve or promote inhibition, respectively [137, 147]. In this way it is a unique example of coupled regulation of both Arf GEF activity and localization independent of phosphoinositides.

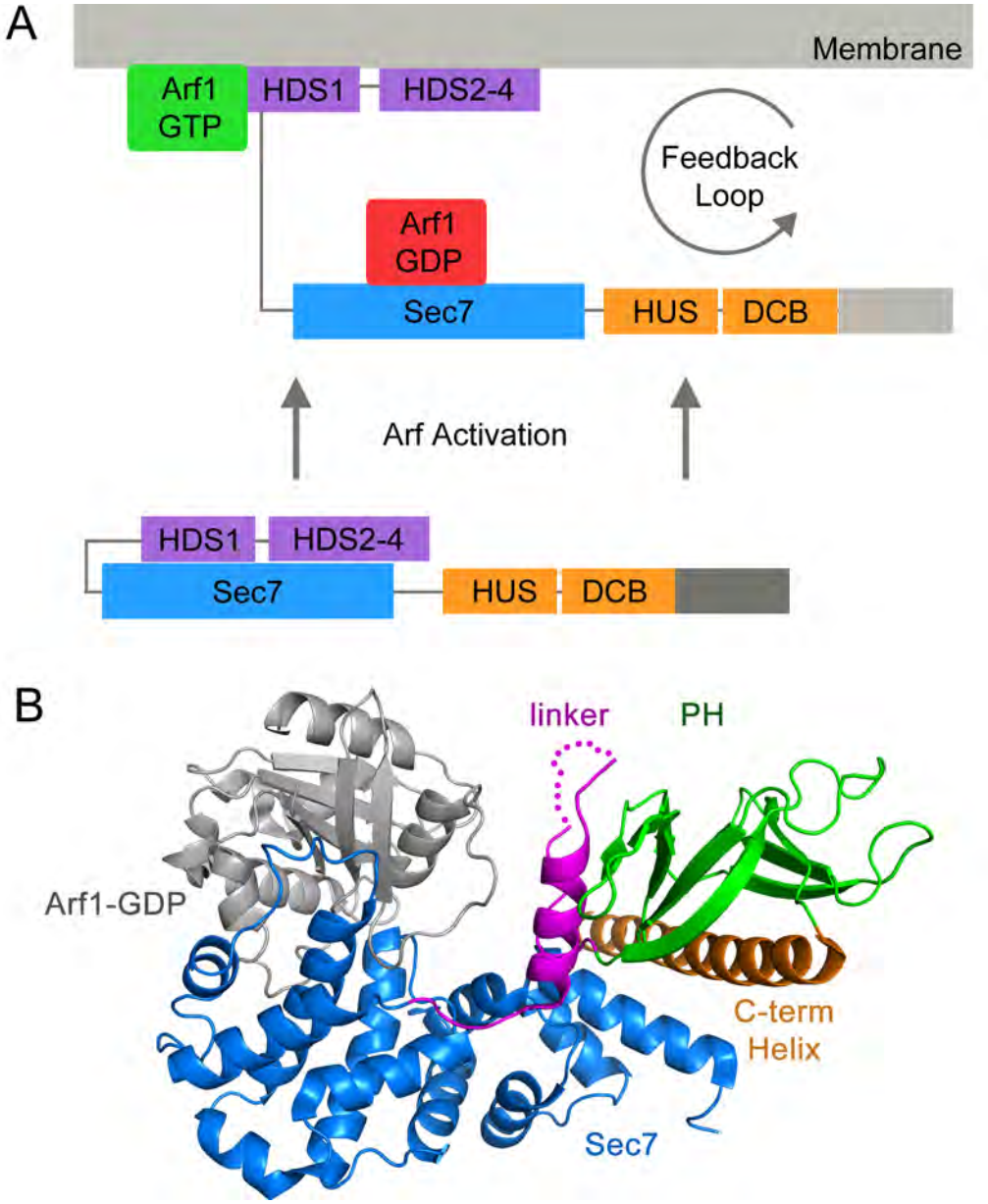

Figure I.13: Autoregulation in Sec7 and BRAG2. A) Schematic depicting relevant interactions in Sec7 autoregulation and a mechanism for autoinhibition relief and membrane recruitment. B) The structure of BRAG2 in complex with Arf1-GDP (PDB ID 4C0A). Unstructured amino acids in the Sec7-PH linker are modeled as circles.

An interesting example of regulation through simple membrane binding elements occurs in BRAG2, a member of the IQSec family of Arf GEFs. It consists of an N-terminal IQ motif which binds calmodulin, proline rich regions, a Sec7 domain, PH domain, and C-terminal coiled coil [78]. BRAG2 shows no clear autoinhibition in solution, and the structure of the protein in the absence of substrate is unknown. However activity in solution is increased by inclusion of the PH domain and enhanced further by membrane binding of either substrate Arf or BRAG2 [87]. Structural studies have found that the Sec7-PH domain linker actively contributes to PH domain stability and that these two regions contact substrate Arf-GDP during catalysis (Figure I.13B). Instead of autoinhibition in this system, the PH domain regulates activity through promiscuous binding of phosphoinositides, which encourages the combined stabilization of the Sec7-PH linker, PH domain, and Arf substrate at membrane environments. This can be seen as an indirect autoregulation in which stabilization of the substrate bound conformation is limited by conformational rearrangements upon membrane binding.

Autoregulation in Cytohesins

In addition to requirements for membrane recruitment in cytohesins discussed above, Cytohesins are structurally autoregulated through physical blockage at the Arf-GDP binding site in the Sec7 domain by substrate mimicry (Figure I.14) [133]. The Sec7-PH domain linker and C-terminal helix tuck hydrophobic residues into the SW1 and SW2 Arf-GDP binding sites, respectively. Key residues Leu 258 and Phe 262 in the linker mimic the packing of Phe 51 in the Arf-GDP SW1 region, and residues Phe 384,

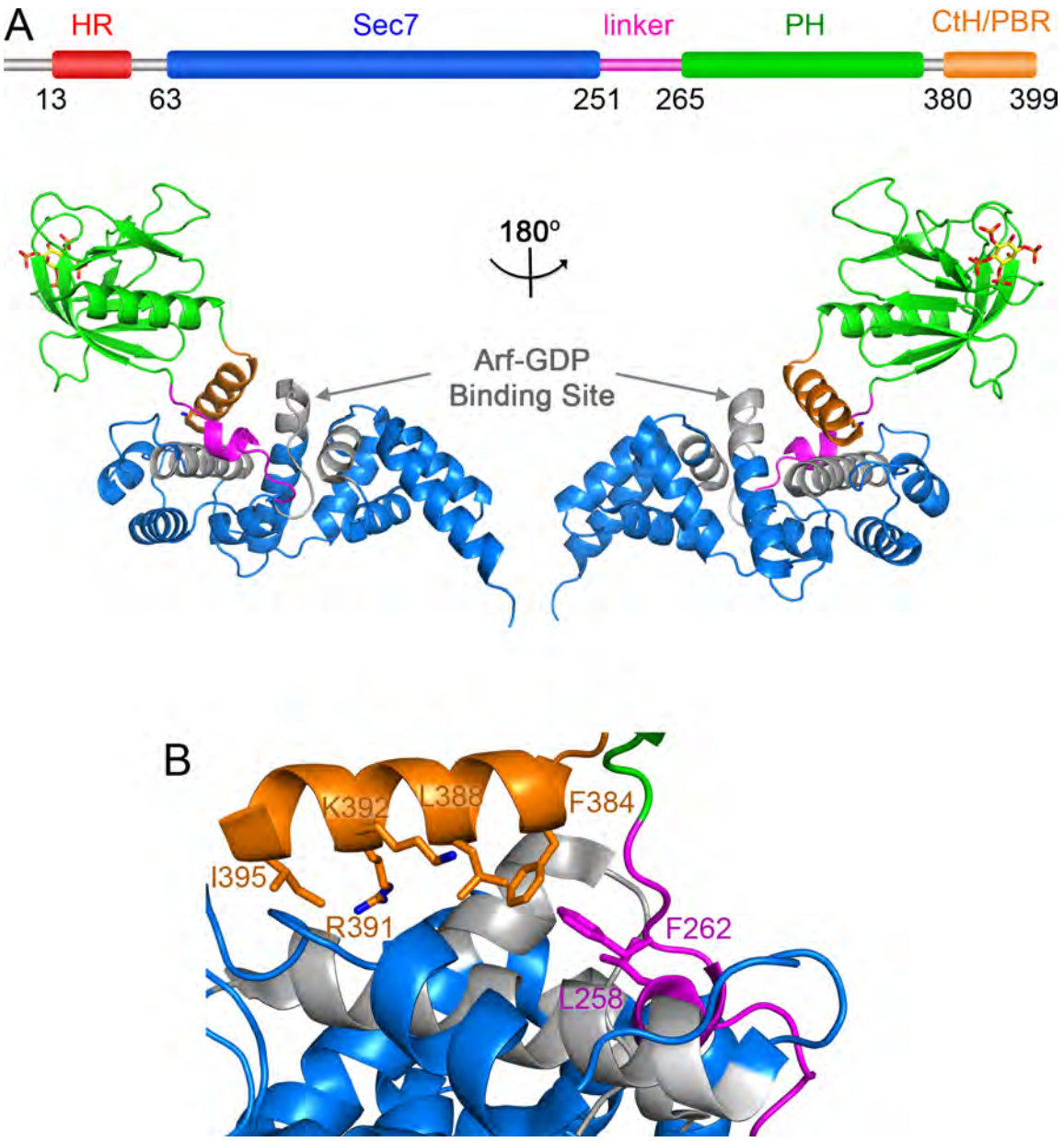

Figure 1.14: The Structure of Autoinhibited Grp1 A) Domain architecture and overall structure of $Grp1_{63\text{-}399}$ (PDB ID 2R09). B) View of autoinhibitory elements docked into the Arf-GDP binding site (gray).

Leu 388, and aliphatic portions of Met 387 and Lys 392 lay over the SW2 docking site. Consistent with the structure, substitution of autoinhibitory residues with alanine resulted in dramatic relief of autoinhibition, with the largest effects coming from L388A and K392A mutations in the CtH. HR dimerization had negligible effects on this process *in vitro*, as both homo and heterodimers showed inhibition that was relieved by truncation of the CtH.

Several studies have linked the molecular mechanism of Cytohesin autoinhibition relief to membrane recruitment, highlighting common themes seen in Sos, Vav1, Sec7, and BRAG2. Most simply, mutation or truncation of residues in the CtH caused reduced plasma membrane recruitment upon insulin stimulation [133]. Phosphorylation of consensus sites for Phosphokinase C (PKC) on the CtH of ARNO and CYTH1 partially relieves autoinhibition and robustly recruits the proteins to membranes in response to phorbol ester [133, 148-150]. Independent of this, Arf6-GTP has been shown to relieve autoinhibition in the presence of the phosphoinositide head group for PIP_3, and recruit Cytohesins to the plasma membrane upon PIP_3 production through interaction with the PH domain [31, 75]. Arf1-GTP relieves autoinhibition in a similar membrane context as Arf6, but the ability of Arf1-GTP to recruit Cytohesins to membrane environments is not fully characterized [31, 151]. Conversely, Arl4 has been shown to recruit Cytohesins to the plasma membrane in a similar manner, but its effect on GEF activity has as of yet not been determined [32, 75].

Thesis Rationale

As described above, much is known about the molecular mechanisms in GTPase biology. For Arf GTPases, structural studies have elucidated the molecular details for nucleotide binding, GDP to GTP exchange, GTP hydrolysis, effector binding, and membrane recruitment [24, 58, 60, 61, 79, 89, 94, 95]. Additionally, Cytohesins have been well characterized with respect to details of GEF catalysis, PH domain phosphoinositide recognition and membrane insertion, and autoregulation [59, 82, 89, 111, 112, 133]. From previous studies it is clear that Arf GTPase and phosphoinositide binding play roles in autoinhibition relief and membrane recruitment of Cytohesins, but the structural basis for this remains unknown. Chapter II of this thesis describes the use of X-ray crystallography, biochemical, and biophysical techniques to define how cytohesins are stimulated and recruited to membrane environments via an integrated structural mechanism.

Generally, active states of autoinhibited molecules have been hard to visualize, presumably due to high degrees of flexibility. Indeed, structural information of all autoinhibited GEFs to date has come from either autoinhibited, or substrate-stabilized conformations, with information on specific regulatory mechanisms deduced from primarily biochemical observations. While these studies have been immensely informative, there are often remaining questions as to conformational flexibility and overall domain organization of these molecules in the active forms. In these cases of potentially flexible systems, Small Angle X-ray Scattering (SAXS) has proven an attractive supplement to crystallography, but obtaining clear, reliable data can prove

challenging. Chapter III highlights some of the challenges of traditional SAXS experiments and documents a novel method of data collection and processing which navigates around them. Additional improvements in data analysis with shape envelopes and modeling algorithms incorporating recent advances in the field will be discussed.

Using the procedures outlined in Chapter III, Chapter IV details a biochemical and SAXS analysis to elucidate the dynamic Sec7 domain orientation in the active Arf-GTP complex. Overall, these findings provide the first structural evidence for a completely integrated mechanism of autoinhibition relief by Arf6-GTP and phosphoinositides that is compatible with structural reorganization of the Sec7 domain.

Chapter II: Structural Basis for Membrane Recruitment and Allosteric Activation in Cytohesin Family Arf GTPase Exchange Factors

Abstract

Membrane recruitment of Cytohesin family Arf guanine nucleotide exchange factors (GEFs) depends on interactions with phosphoinositides and active Arf GTPases, which in turn relieve autoinhibition of the catalytic Sec7 domain through an unknown structural mechanism. Here, we show that Arf6-GTP relieves autoinhibition by binding to an allosteric site that includes the autoinhibitory elements in addition to the PH domain. The crystal structure of a Cytohesin-3 (Grp1) construct encompassing the allosteric site in complex with the head group of phosphatidyl inositol 3,4,5-trisphosphate and N-terminally truncated Arf6-GTP reveals a large conformational rearrangement, whereby autoinhibition can be relieved by competitive sequestration of the autoinhibitory elements in grooves at the Arf6/PH domain interface. Disposition of the known membrane targeting determinants on a common surface is compatible with multivalent membrane docking and subsequent activation of Arf substrates, suggesting a plausible model through which membrane recruitment and allosteric activation could be structurally integrated.

Introduction

Guanine nucleotide exchange factors (GEFs) activate GTPases by catalyzing exchange of GDP for GTP [152]. Since many GEFs are recruited to membranes through interactions with phospholipids, active GTPases, or other membrane-associated proteins [78, 152-155], GTPase activation can be restricted or amplified by spatial-temporal overlap of GEFs with binding partners. GEF activity can also be controlled by autoregulatory mechanisms, which may depend on membrane recruitment [26, 87, 138, 145, 156-158]. Structural relationships between these mechanisms are poorly understood.

Arf GTPases function in trafficking and cytoskeletal dynamics [155, 159, 160]. Membrane partitioning of a myristoylated (myr) N-terminal amphipathic helix primes Arfs for activation by Sec7 domain GEFs [60, 161-163]. Cytohesins comprise a metazoan Arf GEF family that includes the mammalian proteins Cytohesin-1 (Cyth1), ARNO (Cyth2), and Grp1 (Cyth3). The *Drosophila* homolog *steppke* functions in insulin-like growth factor signaling while Cyth1 and Grp1 have been implicated in insulin signaling and Glut4 trafficking, respectively [127, 128, 164]. Cytohesins share a modular architecture consisting of heptad repeats, a Sec7 domain with exchange activity for Arf1 and Arf6, a PH domain that binds phosphatidyl inositol (PI) polyphosphates, and a C-terminal helix (CtH) that overlaps with a polybasic region (PBR) [81, 149, 165-170]. The overlapping CtH and PBR will be referred to as the CtH/PBR. The phosphoinositide specificity of the PH domain is influenced by alternative splicing, which generates di- (2G) and triglycine (3G) variants differing by insertion of a glycine residue in the β1/β2

loop (Figure I.10) [171]. Despite similar PI(4,5)P_2 (PIP_2) affinities, the 2G variant has 30 fold higher affinity for PI(3,4,5)P_3 (PIP_3) [111]. In both cases, PIP_3 is required for plasma membrane (PM) recruitment [149, 166, 172-174], which is promoted by expression of constitutively active Arf6 or Arl4d and impaired by PH domain mutations that disrupt PIP_3 or Arf6 binding, or by CtH/PBR mutations [75, 156, 175, 176].

Cytohesins are autoinhibited by the Sec7-PH linker and CtH/PBR, which obstruct substrate binding (Figure I.14) [156]. Autoinhibition can be relieved by Arf6-GTP binding in the presence of the PIP_3 head group [156]. Active myr-Arf1 and myr-Arf6 also stimulate exchange activity on PIP_2–containing liposomes [177]. Whether this effect is due to relief of autoinhibition per se or enhanced membrane recruitment is not yet clear. Phosphoinositide recognition by PH domains, catalysis of nucleotide exchange by Sec7 domains, and autoinhibition in Cytohesins are well characterized [60, 79, 89, 111, 114, 156, 163, 178-180]. How Arf-GTP binding relieves autoinhibition and promotes membrane recruitment is unknown. In this chapter, we determine the structural basis for relief of autoinhibition and investigate potential mechanistic relationships between allosteric regulation, phosphoinositide binding, and membrane targeting.

Results

Arf6-GTP binding requires autoinhibitory and membrane targeting elements

Though necessary, it is not known if the PH domain is sufficient for Arf6-GTP binding or whether the N-terminal helix of Arf6 contributes to the overall interaction. To delineate the structural requirements, *in vitro* nucleotide exchange and surface plasmon

resonance (SPR) experiments were performed in the presence of Ins(1,3,4,5)P_4 (IP_4) and constitutively active (Q67L) variants of full length Arf6 or a truncated construct lacking the N-terminal helix (Arf6NΔ13). Both Arf6 constructs had similar stimulatory effects on Grp1 exchange activity (Figure II.1), with half maximal constants ($K_{0.5}$ = 18 μM) similar to the value of 14 μM determined in experiments using Arf6-GppNHp [156]. This effect was dependent on IP_4 or PIP_3 on liposomes (Figure II.1C).

To quantify molecular requirements for this interaction, surface plasmon resonance (SPR) experiments were performed in the presence of Ins(1,3,4,5)P_4 (IP_4) using Arf6 and deletion constructs of Grp1. As shown in Figure II.2, Arf6-GppNHp and N-terminally truncated Arf6 Q67L bound $Grp1_{251-399}$, which contains the entire linker and CtH/PBR, with indistinguishable K_D values of 16-18 μM similar to half-maximal activation constants in solution. Comparable dissociation constants were observed for constructs truncating up to 10 residues in the linker region. Affinity was reduced 4 fold by deletion of the entire linker, 2 fold by deletion of two residues at the C-terminus, and was strongly impaired by deletion of the CtH/PBR (K_D > 200 μM). The minimal construct with intact affinity, $Grp1_{260-399}$, includes the CtH/PBR and five residues from the linker.

Structure of the Arf6-GTP complex with Grp1-IP_4

Crystals diffracting to 1.8 Å were obtained for Arf6NΔ13 Q67L bound to $Grp1_{247-399}$ and IP_4. The structure was solved by molecular replacement (Materials and Methods, Table II.1 and Figure II.3). The structures of Arf6-GTP and the PH domain resemble

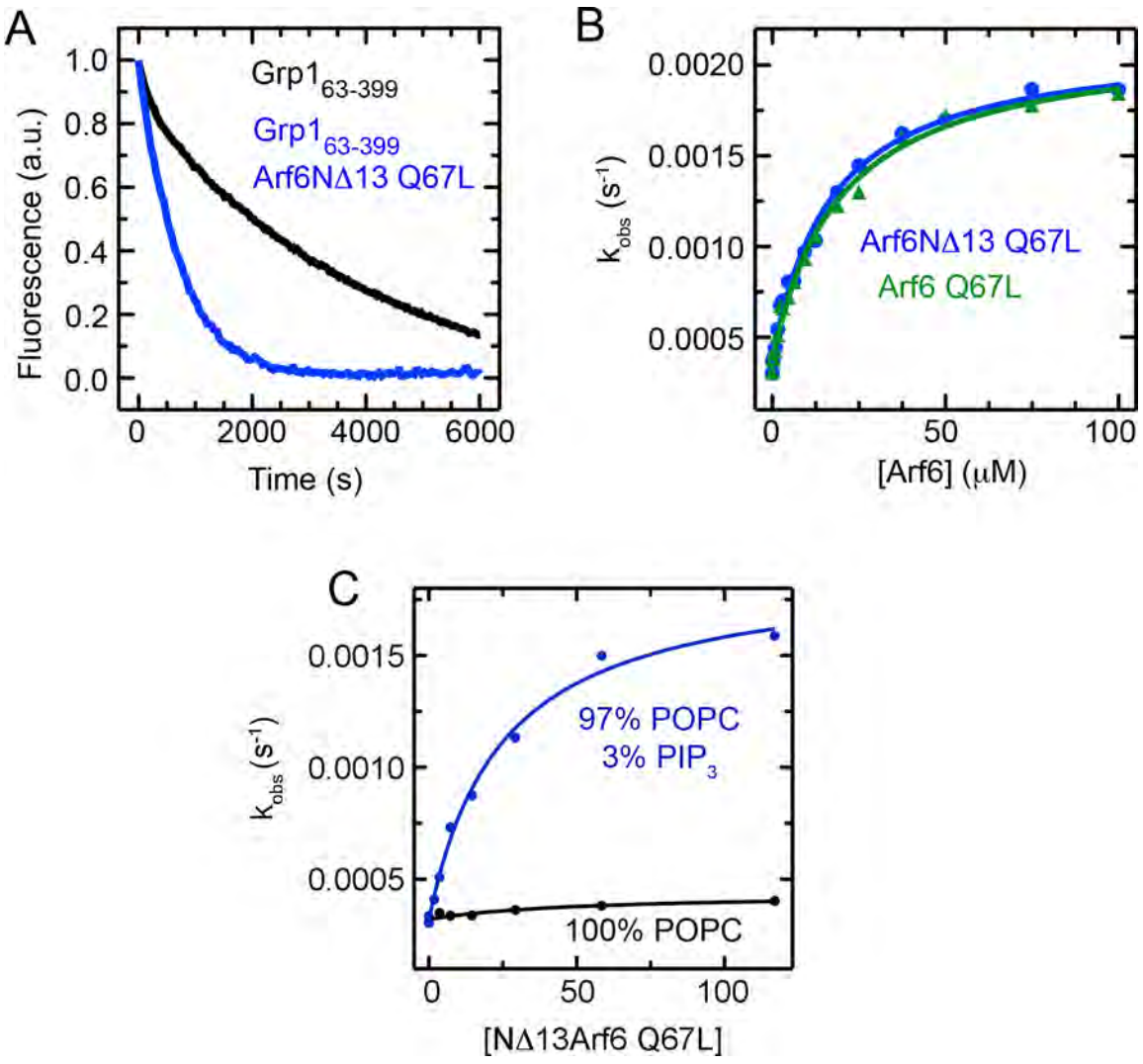

Figure II.1: Relief of Autoinhibition by Arf6-GTP. A) MantGDP dissociation from Arf1NΔ17 catalyzed by 125 nM $Grp1_{63-399}$ in the presence or absence of 80 μM Arf6NΔ13 Q67L. B) Dependence of $Grp1_{63-399}$ catalyzed mantGDP dissociation from Arf1NΔ17 on Arf6 Q67L or Arf6NΔ13 Q67L. C) Dependence of $Grp1_{63-399}$ catalyzed mantGDP dissociation from Arf1NΔ17 on liposome composition in the presence of 80mM Arf6NΔ13 Q67L.

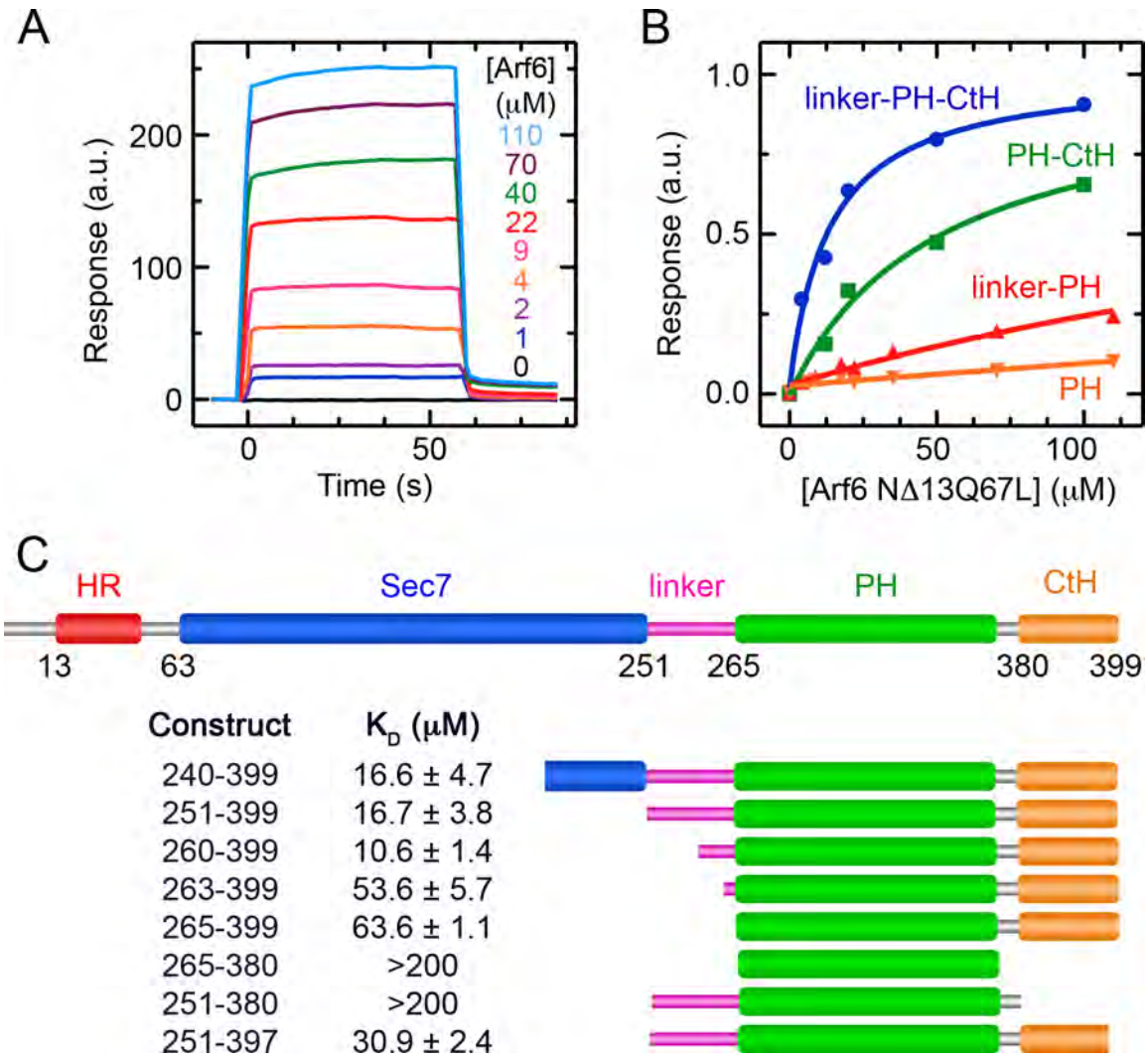

Figure II.2: Direct Binding of Arf6-GTP to Grp1. A) SPR sensorgrams for binding of 6×His Arf6NΔ13 Q67L to immobilized GST-Grp1$_{251-399}$. B) Dependence of equilibrium SPR signals (R_{eq}) for selected Grp1 constructs on Arf6NΔ13 Q67L. C) Dissociation constants for 6×His Arf6NΔ13 Q67L binding to GST-Grp1 constructs determined by SPR. (A-C) Buffers contained 1 μM IP_4.

those of the isolated proteins, as do the interactions with bound ligands [114, 178, 181]. As illustrated in Figure II.3, the complex buries a surface area of 2450 $Å^2$ within a tripartite interface that includes intermolecular interactions with the PH domain as well as inter- and intramolecular interactions involving the linker and CtH/PBR.

At the core of the PH domain interface (Figure II.4A), the β1 strand, switch II, and interswitch regions of Arf6 engage the extended sheet comprised of the β1-β4 strands and βi1-βi2 hairpin insertion in the PH domain. Notably, the βi1-βi2 hairpin is a unique feature of Cytohesins and supplies basic residues critical for phosphoinositide recognition [114, 178]. Invariant Arf6 residues from switch I (Phe 47), interswitch (Trp 62), and switch II (Leu 73, His 76 and Tyr 77) pack against non-polar residues from the β3-β4 strands (Cys 292, Tyr 294, Ile 307 and Pro 309) and βi1-βi2 hairpin (Cys 342 and Val 350) in the PH domain.

At the periphery, interswitch (Asn 48) and switch II (Lys 69) residues mediate polar interactions with, respectively, residues from the β6-βi1 (Lys 340) and β2-β3 (Asp 290) loops in the PH domain. Arg 15 in the β1 strand of Arf6 also donates a hydrogen bond to the main chain carbonyl of Val 350 in the βi2 strand of the PH domain. Despite differences in detail, the PH domain interaction epitope in Arf6 is centered on the invariant hydrophobic triad at the switch/interswitch junction (Figure II.4B) and, in this respect, resembles previously described Arf-effector complexes [62]. The interaction epitope in the PH domain, on the other hand, is distinct from GTPase interaction epitopes in other PH domain complexes, including Arf1-GTP/ARHGAP21 (Figure II.4C) [182].

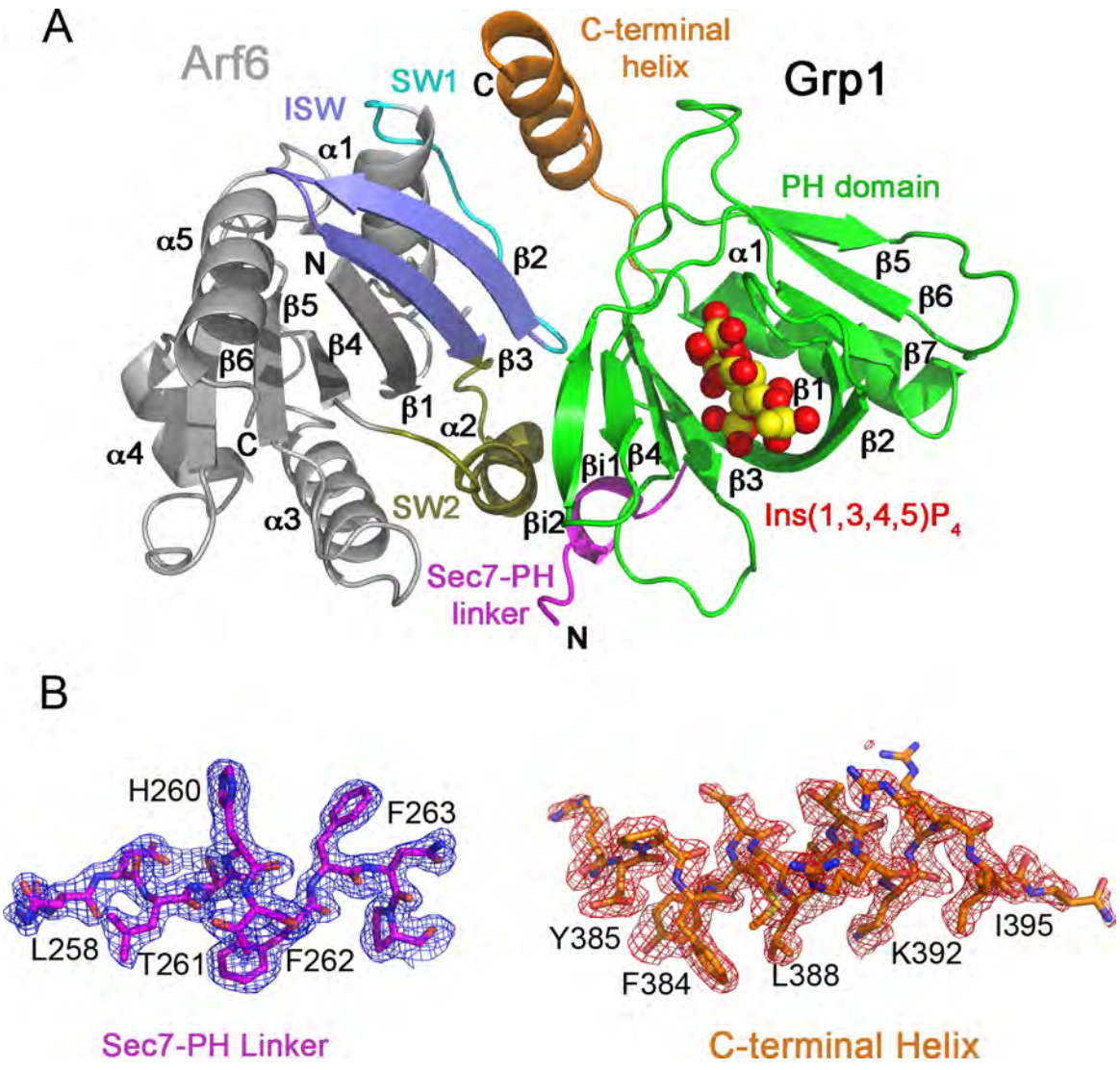

Figure II.3: The Grp1-IP_4:Arf6-GTP Crystal Structure. A) Overall view of the complex with the switch (SW), interswitch (ISW), and other regions colored as indicated. IP_4 is depicted as spheres. B) Omit map for the Sec7-PH linker and C-terminal Helix. The final refined model is overlaid with σA weighted $2F_o$-F_c electron density maps after simulated annealing with the region shown omitted. Maps include data from 39-1.83 Å and are contoured at 1.0 σ.

Arf6NΔ13-GTP / Grp1-Ins(1,3,4,5)P_4 complex	
Data collection	
Space group	$P4_32_12_1$
Cell dimensions	
a, b, c (Å)	56.6, 56.6, 274.4
α, β, γ (°)	90, 90, 90
Resolution (Å)	20–1.83
R_{sym}	0.041 (0.346)
I / σI	24.7 (4.3)
Completeness (%)	98.9
Redundancy	2.7 (2.0)
Refinement	
Resolution (Å)	20–1.83
No. reflections	38345
R_{work} / R_{free}	0.197 / 0.235
No. atoms	
Protein	2501
Ligand/ion	62
Water/solvent	491
B-factors	
Grp1	30.3
Arf6	31.3
Water/solvent	45.2
R.m.s. deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.18

* Data were collected on a single crystal. Values in parentheses are for highest-resolution shell.

Table II.1: Data collection and refinement statistics for the Grp1-IP_4:Arf6-GTP Crystal Structure

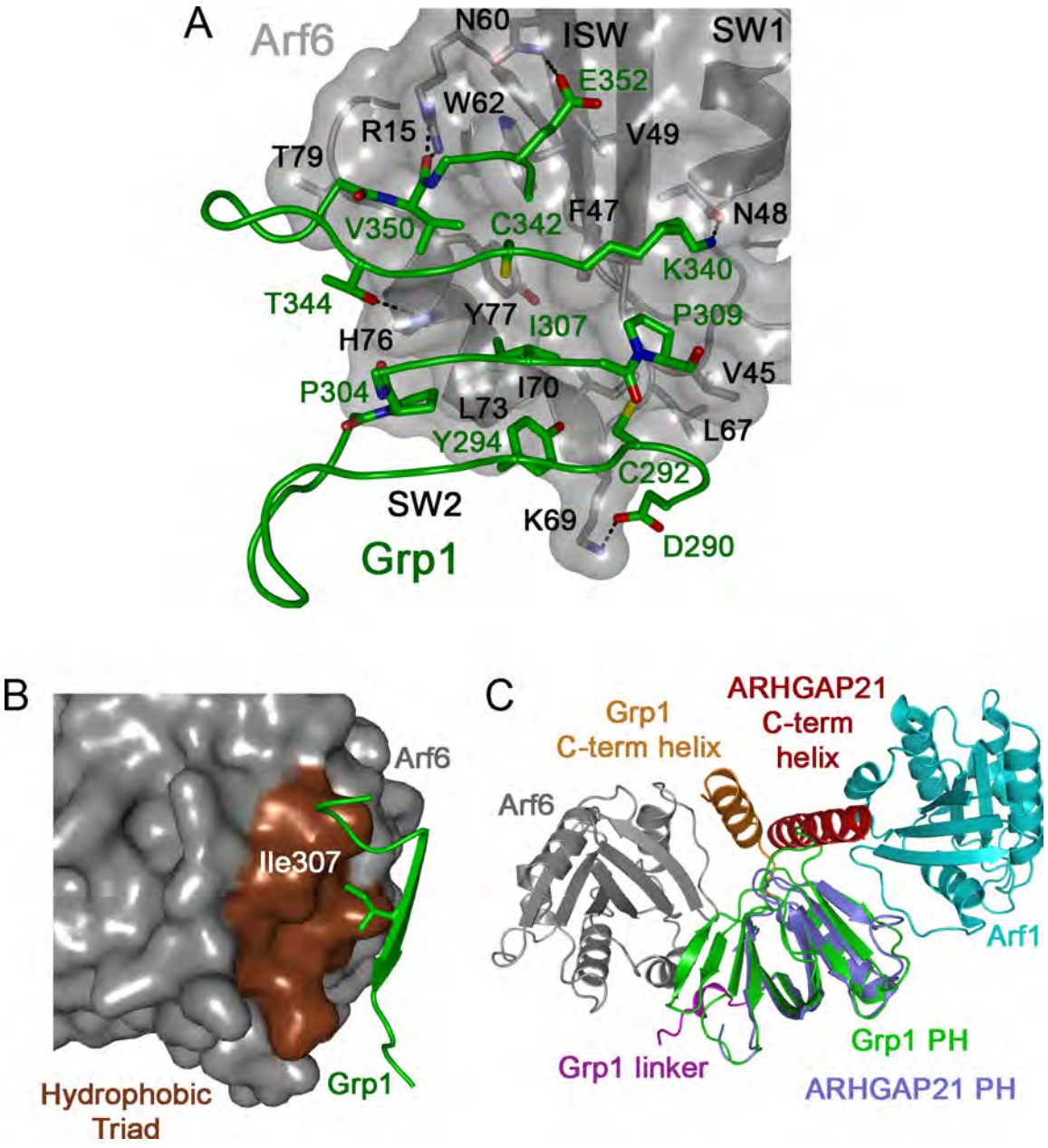

Figure II.4: The Grp1 PH Domain Binding Epitope on Arf6-GTP. A) Core interface between Arf6 and the Grp1 PH domain. B) Grp1 Ile 307 is buried in the canonical non-polar pocket formed by the hydrophobic triad of Arf6. (C) Comparison of the Arf6/Grp1 and Arf1/ARHGAP21 (PDB ID: 2J59) complexes after superposition of their respective PH domains.

The interface of canonical Arf and non-canonical PH domain surfaces facilitates an unusual mode of interaction with the linker and CtH/PBR. The CtH, including the first two residues of the PBR, and the last nine residues in the linker dock in distinct grooves formed at the periphery of the central Arf6-PH domain interface (Figure II.5). In one groove, Leu 258 and Phe 262 in the linker occupy hydrophobic pockets derived from switch II residues (Pro 72, Leu 73, Arg 75, and His 76) and the PH domain β3/β4 loop (Ile 287, Tyr 294, Phe 296, and Lys 307). In the other groove, Phe 384, Leu 388, and Lys 392 from the CtH/PBR pack against a concave surface lined by switch I (Tyr 31) and interswitch (Thr 40 and Ile 42) residues. Whereas the interactions with the CtH/PBR are mediated primarily by Arf6 residues, the linker contacts are more evenly distributed between Arf6 and the PH domain (Figure II.6). These observations explain the stronger reduction in affinity accompanying deletion of the CtH/PBR compared with deletion of the linker, despite similar buried surface areas (757Å^2 vs. 648Å^2, respectively).

Structural basis for relief of autoinhibition by Arf6-GTP

Comparison of the active Arf6-Grp1 and autoinhibited Grp1 structures after superposition of the PH domain reveals large conformational rearrangements in which the linker rotates ~120° about Asp 266 at linker/PH domain junction while the CtH/PBR rotates ~90° about Ser 378 and Ile 379 in the turn preceding the CtH (Figure II.7). The linker rotation is accompanied by changes in its irregular secondary structure. Moreover, several residues that occupy the Sec7 domain exchange site in the autoinhibited structure, including Leu 258 in the linker and Phe 384, Leu 388 and Lys 392 in the CtH/PBR, are

buried in the grooves formed at the Arf6-PH domain interface (Figure II.6 and II.7). Thus, Arf6-GTP binding relieves autoinhibition through an allosteric mechanism involving competitive sequestration of the linker and CtH/PBR in conformations incompatible with autoinhibition.

Determinants of activation and Arf-GTP recognition

The I307E and K340A mutations in the Grp1 PH domain interfere with Arf6-dependent cell spreading and PM recruitment [175, 183]. The equivalent K336A mutation in ARNO impairs myr-Arf6-GTP stimulation of GEF activity for myr-Arf1 on liposomes containing PIP_2 [177]. Lys 340 mediates a polar interaction with Asn 48 in the interswitch region whereas Ile 307 is buried in a pocket formed by the hydrophobic triad (Figure II.4) and conforms to a common theme in Arf:effector recognition [62]. Deletion of the polybasic region in Grp1 or CYTH1 as well as alanine substitution of Leu 388/385 or Lys 392/389 in CYTH disrupted autoinhibition in solution as well as insulin-stimulated PM recruitment [156]. This region makes extensive contact with the hydrophobic groove formed at the Grp1-Arf6 interface. Consistent with this observation, the L388A and K392A substitutions in the CtH/PBR strongly reduce Arf6-GTP binding (Figure II.8).

To further explore the significance of the structural observations and identify determinants for specific activation by Arf6-GTP, residues in the interface were mutated and the effects quantified by measuring direct binding by SPR or catalytic efficiency (k_{cat}/K_M) for Arf1NΔ17-mantGDP/GppNHp exchange in the presence or absence of 80

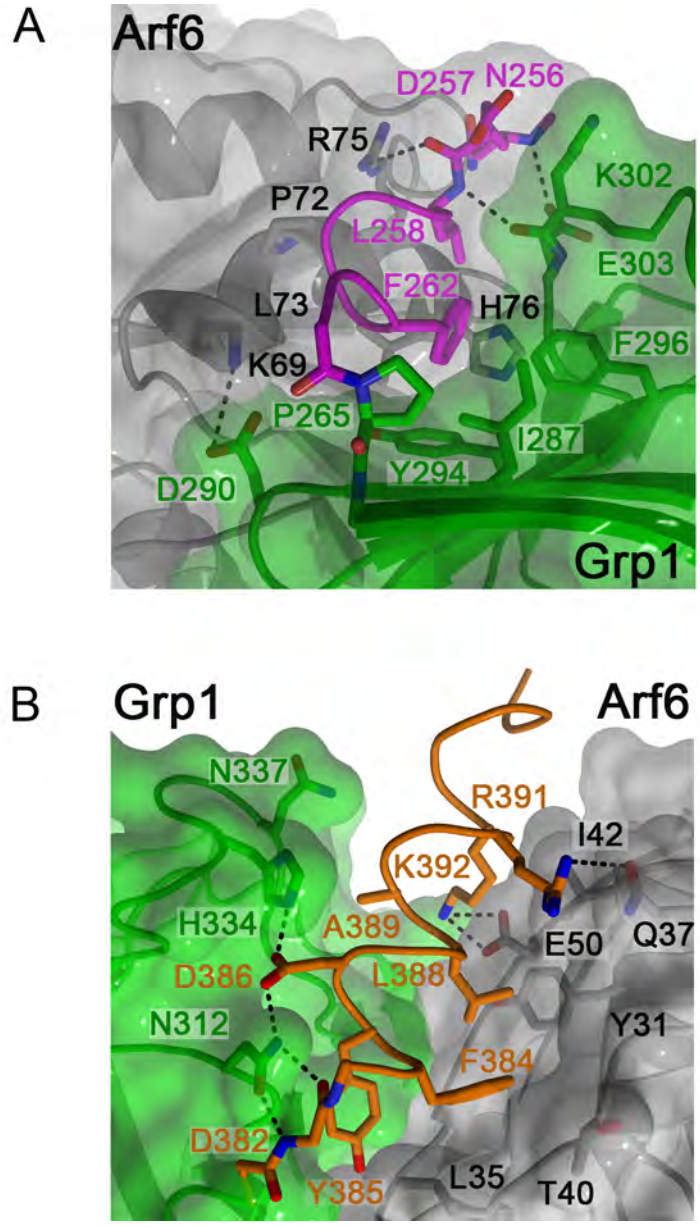

Figure II.5: Interactions of the Grp1 Autoinhibitory Elements within the Grp1-Arf6 Binding Interface. (A and B) Docking of the linker (A) and CtH/PBR (B) within grooves at the periphery of the Arf6-PH domain interface. Molecules and relevant regions are colored as indicated in Figure II.4.

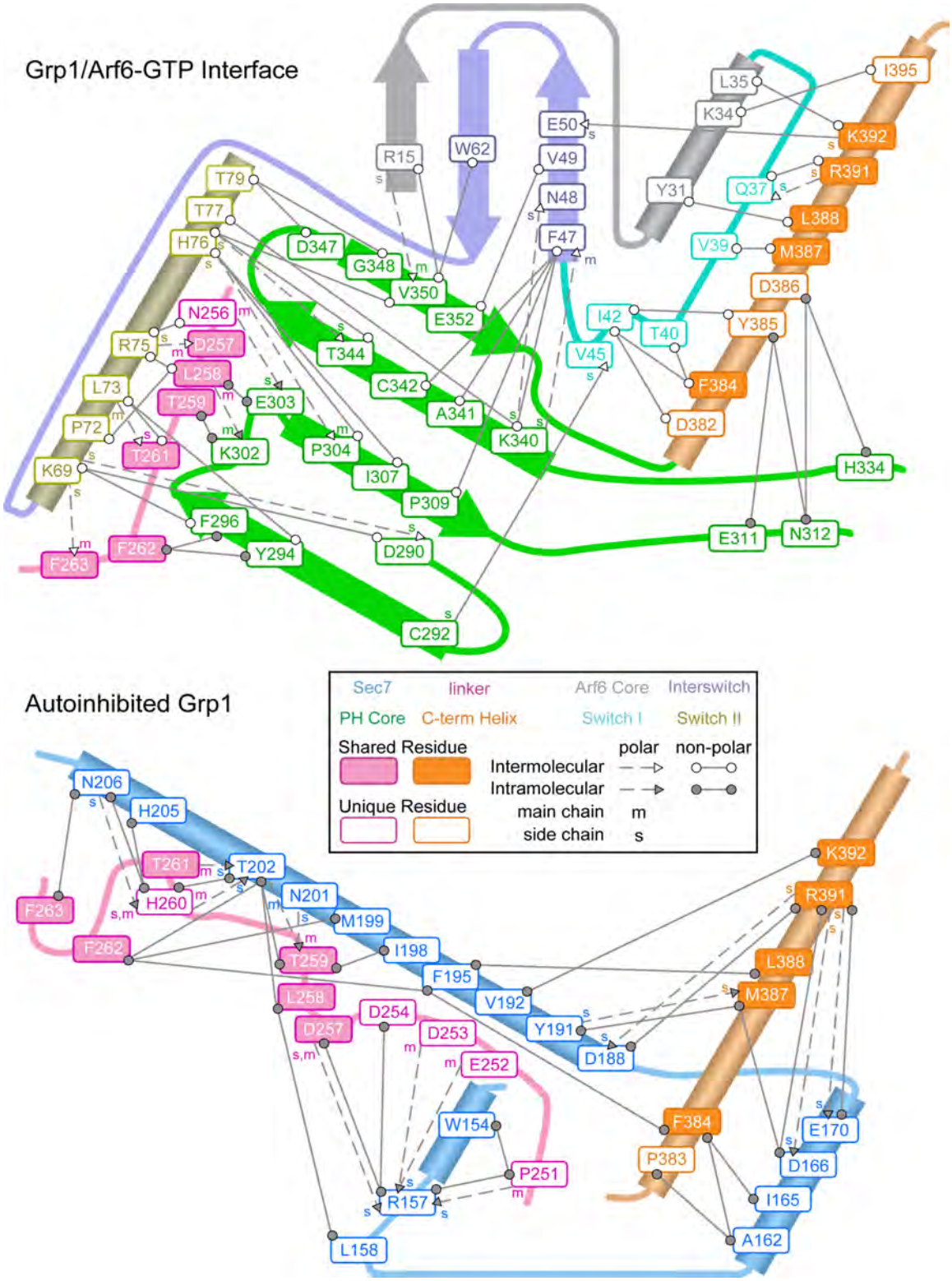
Grp1/Arf6-GTP Interface
Autoinhibited Grp1
Sec7
linker
Arf6 Core
Interswitch
PH Core
C-term Helix
Switch I
Switch II
Shared Residue
Unique Residue
polar
non-polar
Intermolecular
Intramolecular
main chain
side chain
m
s
R15
W62
E50
V49
N48
F47
L35
K34
I395
K392
R391
Q37
Y31
L388
V39
M387
D386
Y385
I42
T40
V45
F384
D382
H334
E311
N312
T79
T77
H76
D347
G348
V350
E352
N256
R75
D257
T344
L258
L73
T259
E303
C342
A341
K340
P72
T261
K302
P304
I307
P309
K69
F296
Y294
D290
F263
F262
C292
N206
H205
T261
T202
F263
H260
N201
M199
F262
T259
I198
F195
L258
V192
D254
D257
Y191
D253
D188
E252
W154
P251
R157
L158
F384
P383
E170
D166
I165
A162

Figure II.6: Detailed Interaction Map of the Grp1-IP_4:Arf6-GTP Complex. Schematic representation of intra- and intermolecular contacts in autoinhibited Grp1 and the active Arf6 complex. Contacts were determined using a 4.0 Å cutoff subject to acceptable stereochemistry.

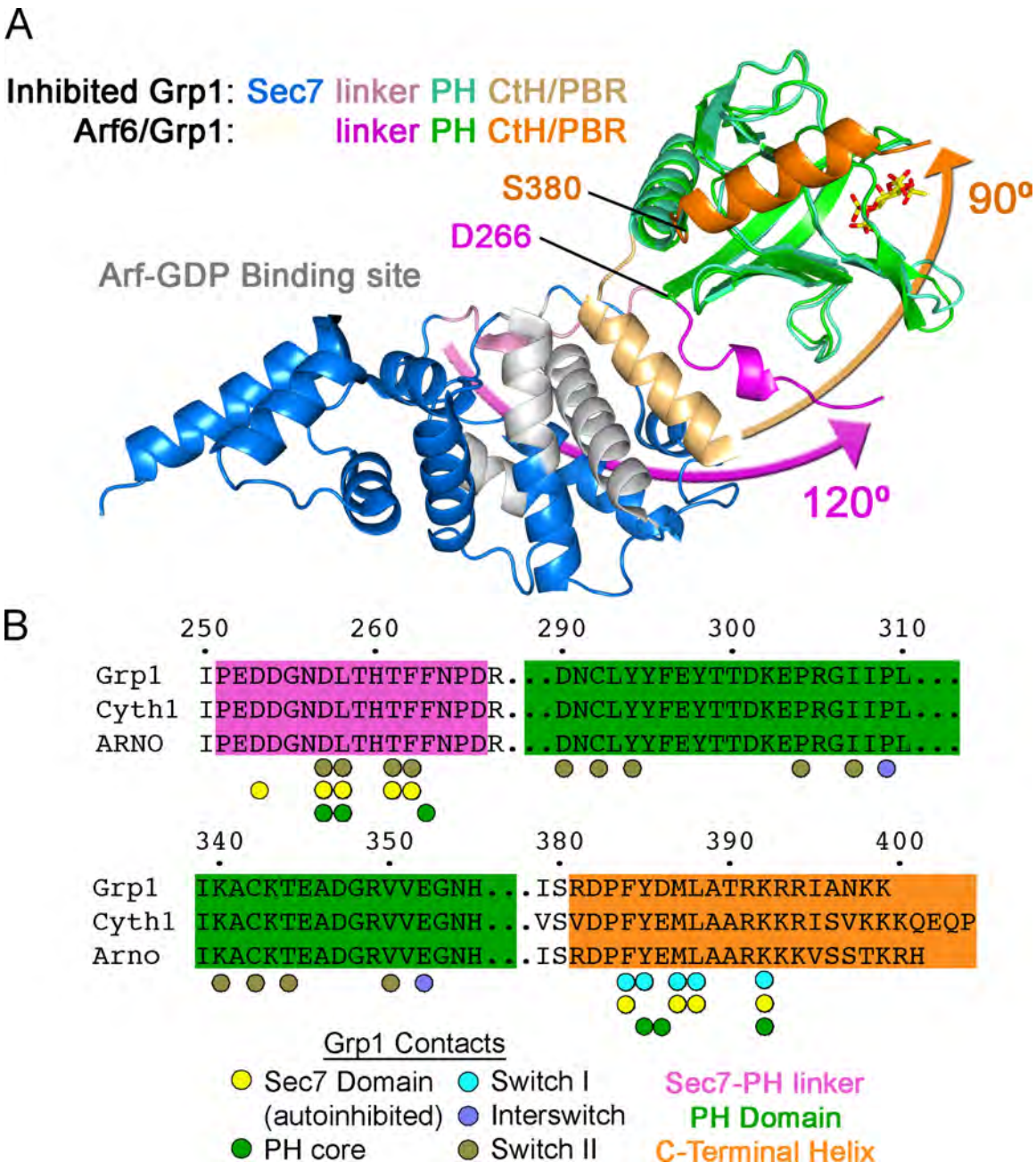

Figure II.7: Competitive Allosteric Mechanism for Relief of Autoinhibition. A) Comparison of autoinhibited Grp1 (PDB ID 2R09) with the active Arf6 complex after superposition of the PH domains. Arrows indicate the angular displacement of the linker and CtH/PBR. B) Sequence Alignment of Grp1 showing shared contacts in autoinhibited and Arf6-bound structures

μM Arf6NΔ13 Q67L. This concentration is 4 fold greater than $K_{0.5}$ for activation and affords robust stimulation while maintaining sensitivity to changes in k_{cat}/K_M. Arf6NΔ13 Q67L increased k_{cat}/K_M 11 fold, which was moderately to severely impaired by switch I (F47A) or switch II (H76A) substitutions (Figure II.8 and Table II.2). PH domain substitutions (Y294A in β3, I307A or I307E in β4, K340A in the β6-βi1 loop, or V350A in βi2) also rendered Grp1 refractory to activation. Finally, alanine substitution of the terminal 2-3 residues in the PBR disrupted autoinhibition and, consequently, activation by Arf6NΔ13 Q67L. Although disordered in both the autoinhibited Grp1 and active Grp1:Arf6 complex structures, the C-terminal lysine residues may contribute to the stability of the CtH through electrostatic interaction with the helix dipole.

To clarify the specificity for active Arf GTPases, relief of autoinhibition by the five human Arf family members was examined (Figure II.8). Compared to Arf6NΔ13 Q67L, stimulation by the N-terminally truncated forms of other Arf GTPases was weaker (2-4 fold). Assuming the extent of activation is similar at saturation, 2-4 fold stimulation is consistent with $K_{0.5}$ values that are 10-25 fold lower than that for Arf6. Given similar tertiary structures, the preference for Arf6 might reflect non-conservative substitutions in the binding interface. Similarity within the interaction epitope (Figure II.8C) suggested five putative specificity determinants. Individual mutation of these residues in Arf1 and Arf6 resulted in reciprocal perturbations of 2 fold or less. The combination of all five substitutions in Arf1 had little additional effect, as was the case for an Arf1→6 chimera that included the complete set of 10 substitutions in the binding epitope (Figure II.8D and Table II.2). Thus, whereas the specificity for some effectors can be reversed by

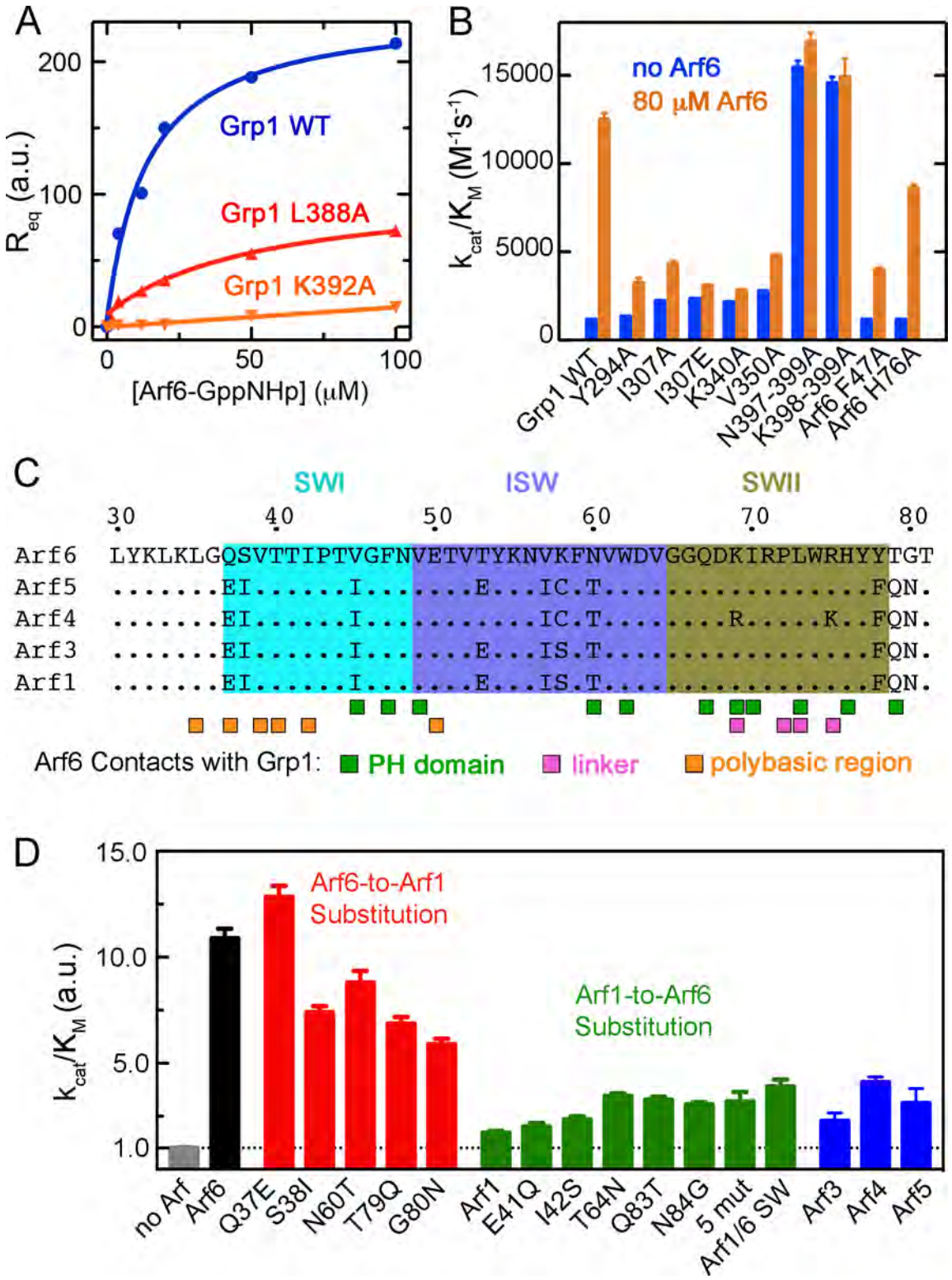
A
200
100
0
R_{eq} (a.u.)
Grp1 WT
Grp1 L388A
Grp1 K392A
0
50
100
[Arf6-GppNHp] (μM)
B
15000
10000
5000
0
k_{cat}/K_M ($M^{-1}s^{-1}$)
no Arf6
80 μM Arf6
Grp1 WT
Y294A
I307A
I307E
K340A
V350A
N397-399A
K398-399A
Arf6 F47A
Arf6 H76A
C
SWI
ISW
SWII
30
40
50
60
70
80
Arf6 LYKLKLGQSVTTIPTVGFNVETVTYKNVKFNVWDVGGQDKIRPLWRHYYTGT
Arf5 .......EI......I.....E...IC.T.....FQN.
Arf4 .......EI......I.....IC.T.....R.....K..FQN.
Arf3 .......EI......I.....E...IS.T.....FQN.
Arf1 .......EI......I.....E...IS.T.....FQN.
Arf6 Contacts with Grp1: PH domain linker polybasic region
D
15.0
10.0
5.0
1.0
k_{cat}/K_M (a.u.)
Arf6-to-Arf1 Substitution
Arf1-to-Arf6 Substitution
no Arf
Arf6
Q37E
S38I
N60T
T79Q
G80N
Arf1
E41Q
I42S
T64N
Q83T
N84G
5 mut
Arf1/6 SW
Arf3
Arf4
Arf5

Figure II.8: Mutational Analysis of the Grp1-IP_4:Arf6-GTP Interaction. (A) Effect of mutations in the CtH/PBR of GST-$Grp1_{251-399}$ on R_{eq} as a function of 6×His Arf6NΔ13-GppNHp in buffer containing 1 μM IP_4. Curves are fitted models. **(**B**)** Effect of mutations on the k_{cat}/K_M of $Grp1_{63-399}$ for Arf1NΔ17 in buffer containing 1 μM IP_4 with or without 80 μM Arf6NΔ13 Q67L. (C) Catalytic efficiency of $Grp1_{63-399}$ in the presence and absence of 80 μM constitutively active, N-terminally truncated Arf GTPases.

Grp1 Construct	Active Arf	k_{cat}/K_M / (10^3 $M^{-1}s^{-1}$)		k_{cat}/K_M (a.u.)*
		GEF	GEF Arf, IP_4	GEF Arf, IP_4
WT	Arf6NΔ13	1.1 ± 0.04	12.5 ± 0.36	11.4 ± 0.33
3G	Arf6NΔ13	1.8 ± 0.13	3.8 ± 0.31	2.1 ± 0.17
Y294A	Arf6 NΔ13	1.3 ± 0.03	3.2 ± 0.29	2.5 ± 0.22
I307A	Arf6NΔ13	2.2 ± 0.02	4.3 ± 0.18	2.0 ± 0.08
I307E	Arf6NΔ13	2.3 ± 0.04	3.1 ± 0.07	1.3 ± 0.03
K340A	Arf6NΔ13	2.1 ± 0.06	3.6 ± 0.01	1.7 ± 0.005
V350A	Arf6NΔ13	2.8 ± 0.05	4.8 ± 0.11	1.7 ± 0.06
397NKK/AAA399	Arf6NΔ13	15.4 ± 0.39	16.9 ± 0.52	1.1 ± .03
K398A/K399A	Arf6NΔ13	14.6 ± 0.36	14.9 ± 1.0	0.97 ± 0.06
WT	Arf6NΔ13 Q37E		14.7 ± 0.40	13.4 ± 0.36
WT	Arf6NΔ13 S38I		8.5 ± 0.18	7.7 ± 0.16
WT	Arf6NΔ13 F47A		4.0 ± 0.14	3.6 ± 0.13
WT	Arf6NΔ13 N60T		10.1 ± 0.56	9.2 ± 0.51
WT	Arf6NΔ13 T79Q		7.9 ± 0.26	7.2 ± 0.24
WT	Arf6NΔ13 G80N		6.8 ± 0.21	6.2 ± 0.19
WT	Arf1NΔ17-GTP		1.9 ± 0.08	1.7 ± 0.09
WT	Arf1NΔ17 E41Q-GTP		2.3 ± 0.17	2.0 ± 0.16
WT	Arf1NΔ17 I42S-GTP		2.7 ± 0.14	2.3 ± 0.15
WT	Arf1NΔ17 T64N		3.6 ± 0.33	3.3 ± 0.3
WT	Arf1NΔ17 Q83T		4.0 ± 0.07	3.6 ± 0.06
WT	Arf1NΔ17 N84G		3.8 ± 0.10	3.5 ± 0.09
WT	Arf1NΔ17 EITQN/QSNTG		3.7 ± 0.51	3.4 ± 0.46
WT	Arf1NΔ17 to Arf6 SW		4.5 ± 0.35	4.1 ± 0.32
WT	Arf3NΔ14		2.6 ± 0.43	2.4 ± 0.39
WT	Arf4NΔ17-GTP		4.7 ± 0.26	4.1 ± 0.26
WT	Arf5ND14		3.6 ± 0.78	3.3 ± 0.71

Table II.2: Catalytic efficiencies for $Grp1_{63-399}$ with active Arf GTPases. All experiments with Arf and IP4 were at concentrations of 80 μM and 1 μM, respectively. (*) Relative to the basal k_{cat}/K_M for each WT or mutant Grp1 protein.

interfacial substitutions in the switch/interswitch regions [184], the interaction with Grp1 is also sensitive to determinants outside the binding epitope.

Relationship between phosphoinositide and Arf6-GTP binding

The IP_4 requirement is a striking property of Arf6-GTP binding and relief of autoinhibition [156, 175]. The underlying structural basis may be related to conformational changes in the βi1/βi2 hairpin and β3/β4 loop, which mediate key contacts with both the head group and Arf6-GTP. Indeed, subtle structural differences in these elements have been described for IP_4 binding to the Grp1 PH domain [114]. Head group binding may also overcome an electrostatic barrier resulting from proximity of the electropositive PBR and head group site in the Arf6-GTP complex.

Substantially higher Arf6-GTP concentrations are required to stimulate 3G ARNO or 3G Cyth1 compared to the 2G variants [156]. Since the residues in the Arf6-GTP binding epitope are conserved and the β1/β2 loop containing the splice site is located distal to the interface (Figure II.9A), the differences in stimulation are likely an indirect effect of reduced affinity of the 3G variants for IP_4 [111]. Consistent with this prediction, the 2G variants bound Arf6NΔ13 Q67L with comparable affinity and were robustly stimulated by Arf6NΔ13 Q67L (Figure II.9B and C). The 3G variants, on the other hand, exhibited muted stimulation that was enhanced by increasing the IP_4 concentration. Similar results were obtained for the T280A mutation in the β1/β2 loop, which eliminates polar interactions with the 1-phosphate and reduces the affinity for IP_4

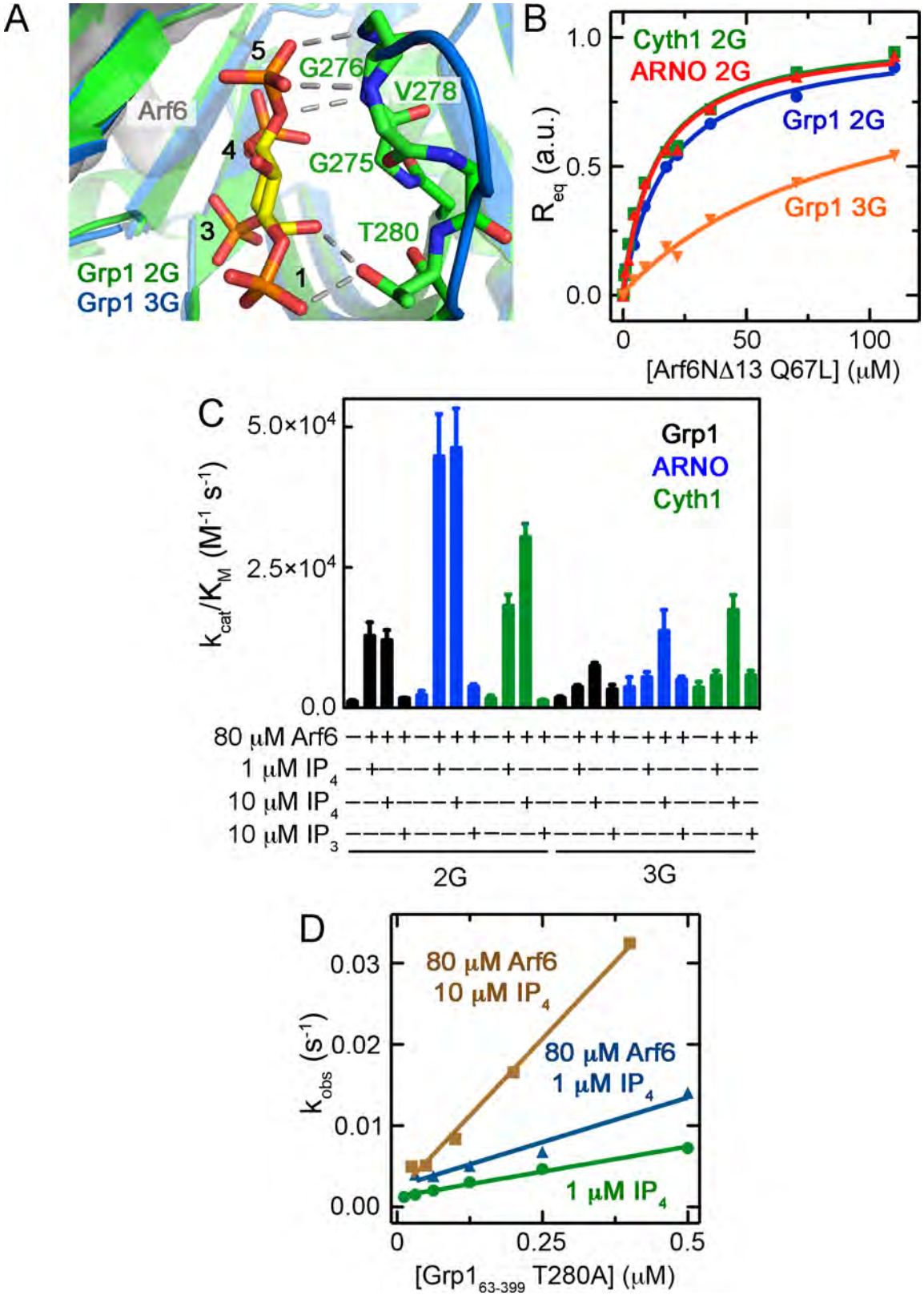
A
Arf6
5
G276
V278
4
G275
3
T280
1
Grp1 2G
Grp1 3G
B
Cyth1 2G
ARNO 2G
Grp1 2G
Grp1 3G
R_{eq} (a.u.)
1.0
0.5
0.0
0
50
100
[Arf6NΔ13 Q67L] (μM)
C
k_{cat}/K_M (M^{-1} s^{-1})
5.0×10^4
2.5×10^4
0.0
Grp1
ARNO
Cyth1
80 μM Arf6 −+++−+++−+++−+++−+++−+++
1 μM IP_4 −+−−−+−−−+−−−+−−−+−−−+−−
10 μM IP_4 −−+−−−+−−−+−−−+−−−+−−−+−
10 μM IP_3 −−−+−−−+−−−+−−−+−−−+−−−+
2G
3G
D
k_{obs} (s^{-1})
0.03
0.02
0.01
0.00
80 μM Arf6
10 μM IP_4
80 μM Arf6
1 μM IP_4
1 μM IP_4
0
0.25
0.5
[$Grp1_{63-399}$ T280A] (μM)

Figure II.9: Effect of Cytohesin Splice Variants on Arf6-GTP Binding and Nucleotide Exchange. A) Comparison of head group binding sites in the 2G Grp1 PH domain from the Arf6 complex and the 3G Grp1 PH domain bound to IP_4 (PDB ID 1U2B) after superposition. B) R_{eq} for 6×His Arf6NΔ13 Q67L binding to 2G GST-$Grp1_{251-399}$, 3G GST-$Grp1_{251-399}$, 2G $ARNO_{247-397}$, and 2G GST-$Cyth1_{251-399}$. Curves are fitted models. C) Effect of 2G vs. 3G splice variation on the k_{cat}/K_M of $Grp1_{63-399/400}$, $ARNO_{57-397/398}$, and $Cyth1_{57-399/400}$ for Arf1NΔ17-mantGDP in the presence or absence of IP_4, IP_3, and/or Arf6NΔ13 Q67L. D) Effect of IP_4 on catalysis of mantGDP release from Arf1NΔ17 as a function of $Grp1_{63-399}$ T280A. Lines are linear fits.

Cytohesin	k_{cat}/K_M / (10^3 $M^{-1}s^{-1}$)			
	GEF	GEF + NΔ13Arf6 Q67L		
		1 μM IP_4	10 μM IP_4	10 μM IP_3
Grp1 2G T280D	1.3 ± 0.07	2.1 ± 0.38	7.0 ± 0.45	---
Grp1 2G	1.1 ± 0.04	12.5 ± 0.36	12.0 ± 1.8	1.8 ± 0.03
Grp1 3G	1.8 ± 0.13	3.8 ± 0.31	7.4 ± 0.63	3.3 ± 0.83
ARNO 2G	1.8 ± 0.27	44.7 ± 7.5	46.3 ± 7.0	3.8 ± 0.04
ARNO 3G	3.7 ± 1.7	5.4 ± .96	13.7 ± 3.7	5.0 ± 0.53
Cyth1 2G	1.7 ± 0.56	18.2 ± 2.0	30.4 ± 2.4	1.2 ± 0.20
Cyth1 3G	3.6 ± 1.0	5.7 ± .94	17.5 ± 2.6	5.8 ± 0.84

Cytohesin	k_{cat}/K_M (a.u.) *		
	GEF + NΔ13Arf6 Q67L		
	1 μM IP_4	10 μM IP_4	10 μM IP_3
Grp1 2G T280D	1.6 ± 0.29	5.4 ± 0.35	---
Grp1 2G	11.4 ± 0.33	10.9 ± 1.6	1.6 ± 0.03
Grp1 3G	2.1 ± 0.17	4.1 ± 0.35	1.8 ± 0.46
ARNO 2G	24.8 ± 4.2	25.7 ± 3.9	2.1 ± 0.02
ARNO 3G	1.5 ± 0.26	3.7 ± 1.0	1.4 ± 0.15
Cyth1 2G	10.7 ± 1.2	17.9 ± 1.4	0.71 ± 0.12
Cyth1 3G	1.6 ± 0.26	4.8 ± 0.72	1.6 ± 0.23

Table II.3: Catalytic Efficiencies of Cytohesin Variants. (*) Relative to the basal k_{cat}/K_M for each Cytohesin protein

by 13 fold (Figure II.9d) [111]. Replacing IP_4 with Ins(1,4,5)P_3 (IP_3) also substantially diminished Arf6NΔ13 Q67L stimulation (Figure II.9, Table II.3).

Relief of autoinhibition in a membrane environment

Membrane partitioning of Cytohesins and myristoylated Arfs results in high local concentrations within the restricted volume proximal to the membrane surface. Since binding of myr-Arf-GTP can enhance membrane recruitment and since the product of the exchange reaction is also a potential activator/recruiter, experiments analyzing relief of autoinhibition on membranes must distinguish allosteric activation from restricted volume effects and feedback amplification [177]. Thus, Arf1NΔ17-mantGDP was used as a non-partitioning substrate reporter for the autoinhibitory status of 2G Grp1 in the presence and absence of myr-Arf6 Q67L or Arf6NΔ13 Q67L and liposomes with or without PIP_3 or PIP_2 (Figure II.10A). Under the conditions of these experiments, the fraction of Grp1 partitioned with membranes in the absence of myr-Arf6 Q67L is >60% for LUVs containing PIP_3 and <10% for LUVs containing either 3% PIP_2 or no phosphoinositide (Figure II.10B). As shown in Figure II.11, the exchange activity of Grp1 increased substantially with increasing concentrations of myr-Arf6 Q67L in the presence of liposomes containing PIP_3 ($K_{0.5} \approx 75$ nM) whereas weaker stimulation was observed in the presence of liposomes containing PIP_2 ($K_{0.5} \approx 350$ nM) or no phosphoinositide ($K_{0.5} \approx 800$ nM). Truncation of the Arf6 N-terminus increased $K_{0.5}$ by 2-3 orders of magnitude whereas the K340A mutation in the PH domain, which disrupts

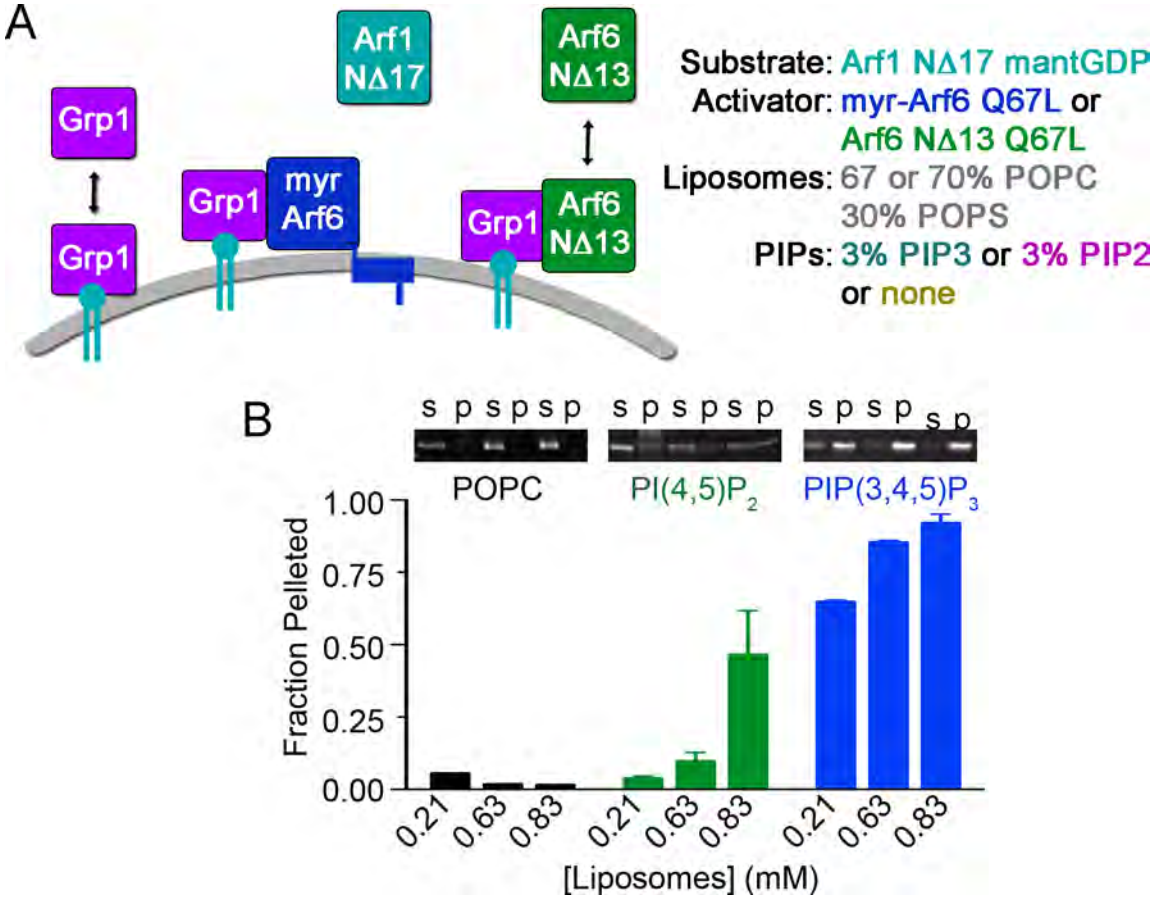

Figure II.10: Experimental Setup for Studying Arf6 Stimulation of Cytohesin Activity in Membrane Environments Using Soluble Substrate. A) Schematic of protein/lipid interactions and compositions relevant to Figure II.11. B) Partitioning of $Grp1_{63-399}$ with liposomes. 500 nM $Grp1_{63-399}$ was incubated with liposomes containing 100% POPC, 97% POPC:3% PtdIns(4,5)P_2, or 97% POPC:3% PtdIns(3,4,5)P_3 for 15 min at 22°C. Reactions were spun at 100,000 × g for 30 minutes and the supernatant and pellets ran on an SDS-PAGE gel stained with Krypton™ stain (Pierce). Gels were imaged with a Kodak Image Station 4000mm using excitation and emission wavelengths of 520-530 and 580-600 nm, respectively. Bands were quantified by gel densitometry using ImageJ.

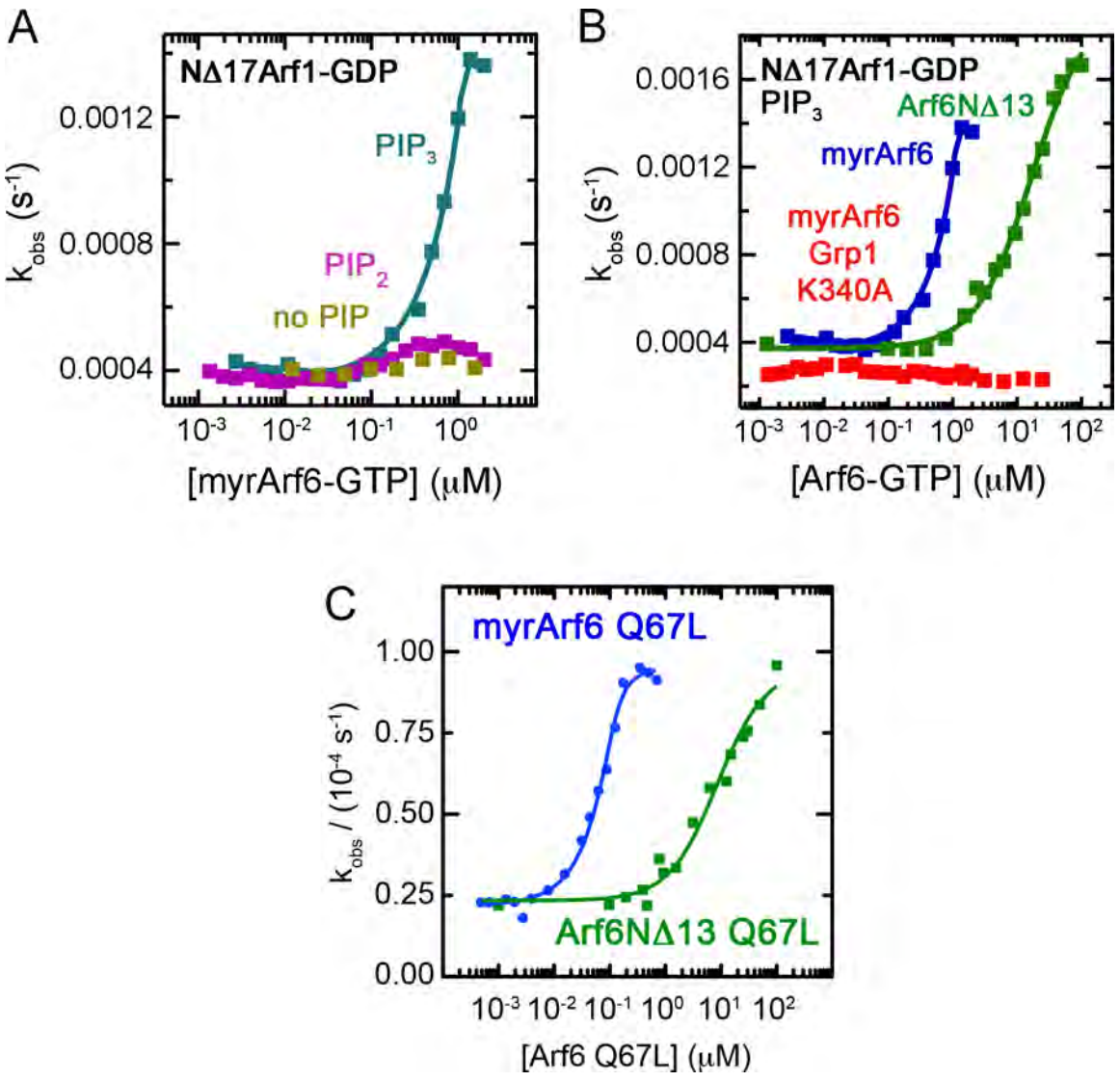

Figure II.11: Atf6-Stimulation of Cytohesin Activity in Membrane Environments Using Soluble Substrate. A) Catalysis of mantGDP release from 1 μM Arf1NΔ17 by 125 nM $Grp1_{63-399}$ as a function of myr-Arf6 Q67L-GppNHp. Curves are fitted quadratic models. B**)** Catalysis of mantGDP release from 1 μM Arf1NΔ17 by 125 nM $Grp1_{63-399}$ or the K340A mutant as a function of myr-Arf6 Q67L-GppNHp or Arf6NΔ13 Q67L-GTP. C) Catalysis of mantGDP release from 1 μM Arf1NΔ17 by 125 nM $Arno_{58-399}$ as a function of myr-Arf6 Q67L-GppNHp or Arf6NΔ13 Q67L-GTP. Experiments were performed in the presence of 0.2 mM LUVs containing 67% POPC:30% POPS:3% PtdIns(3,4,5)P_3.

Arf6-GTP binding, eliminated the stimulatory effect. Similar stimulation was observed for 2G ARNO in the presence of liposomes containing PIP_3 (Figure II.11).

When experiments were performed with myr-Arf1-mantGDP as the substrate and myr-Arf6 Q67L as the activating GTPase (Figure II.12A), strong stimulation was observed in the presence of PIP_3 ($K_{0.5} \approx 50$ nM), moderate stimulation in the presence of PIP_2 ($K_{0.5} \approx 170$ nM), and weaker stimulation in the absence of phosphoinositide ($K_{0.5} \approx 330$ nM) (Figure II.12B). Substantially weaker stimulation in the presence PIP_3 was observed for myr-Arf1 Q71L ($K_{0.5} \approx 1200$ nM) as well as the Grp1 I307E mutant ($K_{0.5} \approx 1600$ nM) (Figure II.12C). Although the experiments combining myristoylated substrate and activator GTPases do not distinguish membrane recruitment from allosteric activation per se, the overall trends are consistent with both the experiments using truncated substrate and myristoylated activator (Figure II.11) as well as experiments using truncated substrate and activator GTPases in the absence of membranes (Figure II.8). This indicates that biochemical preferences for the protein-protein and protein-lipid interactions cooperate to achieve robust stimulation of GEF activity, the efficiency of which is highly dependent on membrane association.

Discussion

We have shown that high affinity Arf6-GTP binding to Grp1 requires the proximal elements previously implicated in autoinhibition and/or membrane targeting. The structure of the complex revealed an unusual binding modality whereby formation of the Arf6-GTP/PH domain interface creates grooves that sequester the linker and

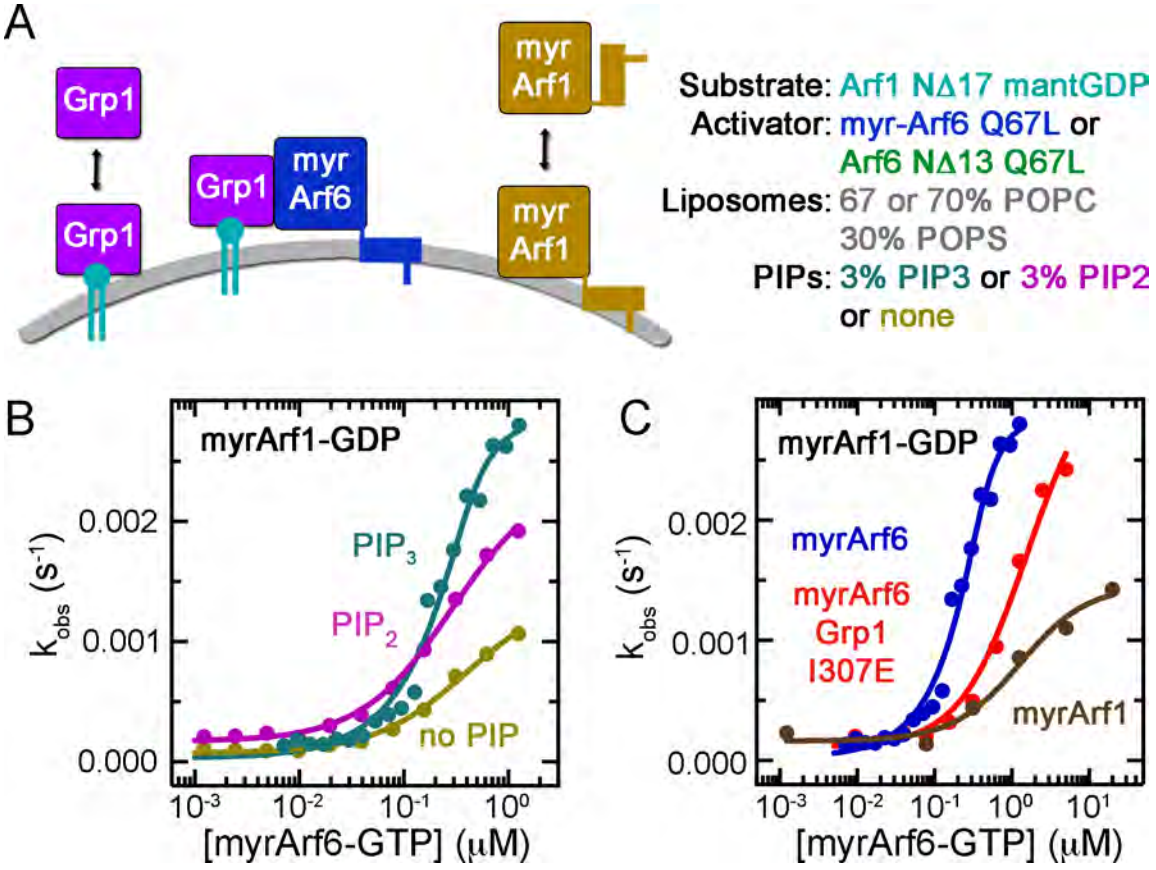

Figure II.12: Arf6-Stimulation of Cytohesin Activity in Membrane Environments Using Membrane Associated Substrate. A) Schematic of protein/lipid interactions and liposome compositions relevant to B) and C). B) Catalysis of mantGDP release from 1 μM myr-Arf1 by 10 nM $Grp1_{63-399}$ as a function of myr-Arf6 Q67L-GTP. Curves are fitted quadratic (PIP_3) or hyperbolic models (PIP_2 and no PIP). C) Catalysis of mantGDP release from 1 μM myr-Arf1 by 10 nM $Grp1_{63-399}$ or the I307E mutant as a function of myr-Arf6 Q67L-GTP. Curves are fitted quadratic models.

CtH/PBR. Substitutions within the binding interface contribute substantially to the observed specificity, but do not fully explain the preference for Arf6. Unlike Arf6 specific effectors studied to date, Grp1 engages common Arf recognition determinants [62] attributed to Arf1-GTP binding and has a more expansive interface, which is shifted away from the interswitch region compared to the Arf6-specific effector JIP4 (Figure I.7) [184] and is thus less sensitive to interswitch substitutions. Careful experimentation will be required to delineate the remaining specificity determinants, which might include substitutions proximal to the binding interface and/or differences in isoelectric point. Moreover, since Arf6 is expressed at lower levels than Arf1 and has a steady state localization at the PM and on endosomes rather than the Golgi [185], the higher affinity for Arf6-GTP may have evolved to compensate for lower Arf6 levels. Consistent with the observation that co-expression of ARNO with either Arf1 or Arf6 promotes Arf1 recruitment to the PM [175], activation by Arf6-GTP may initiate a feedback amplification loop whereby Arf-GTP products further stimulate GEF activity [177].

The diminished ability of Arf6-GTP to stimulate 3G ARNO or 3G Cyth1 compared to 2G Grp1 stems entirely from the splice variation, which reduces PIP_3 affinity and, consequently, the available pool of phosphoinositide-primed precursor. Since the effective concentrations of membrane-associated Cytohesins and myristoylated Arf GTPases may exceed those in the cytoplasm by several orders of magnitude, weak stimulation in solution does not exclude more robust stimulation on membranes, as observed for myr-Arf1-GTP and 3G ARNO on liposomes containing PIP_2 [177]. Notably, however, relief of autoinhibition shows similar characteristics in solution and

membrane environments. Indeed, the main difference lies in the concentration range of activator GTPase required to stimulate the exchange activity. CYTH1 and ARNO are expressed predominantly in the 3G form and Grp1 and Cytohesin-4 in the 2G form [126]. This coupled with the fact that Cytohesin-4 is expressed almost exclusively in blood leukocytes [126] suggest that splice differences may have evolved for attenuated or specialized signaling characteristics from these isoforms in varied cell types and environments.

Taken together, the structural and biochemical data suggest a plausible model integrating membrane targeting with allosteric activation. As depicted in Figure II.13, the 1-phosphate of the bound IP_4 can be aligned with the 1-phosphates of a simulated bilayer [186] in an orientation optimal for simultaneous partitioning of elements known to insert into the hydrocarbon core – in particular the myristoylated N-terminal helix of Arf6-GTP and residues in the β1/β2 loop (Val 278), β3/β4 loop (Tyr 298) and βi1/βi2 hairpin (Ala 346) of the PH domain [60, 124, 163]. Other elements implicated in electrostatic interactions with anionic phospholipids, including the PBR and a 'sentry glutamate' that reduces basal membrane association [119, 170, 173], are located proximal to the membrane surface as are two basic residues in the β5/β6 loop (Arg 322 and Lys 323). Lateral extension of the linker suggests that the Sec7 domain may also be closer to the membrane in the Arf6-GTP complex than in the autoinhibited conformation.

Autoinhibition has been observed in structurally unrelated GEFs for different GTPase families and can be relieved by various inputs, including phosphorylation and interaction with proteins or phospholipids [136, 138, 143, 145, 158]. In Sos, for example,

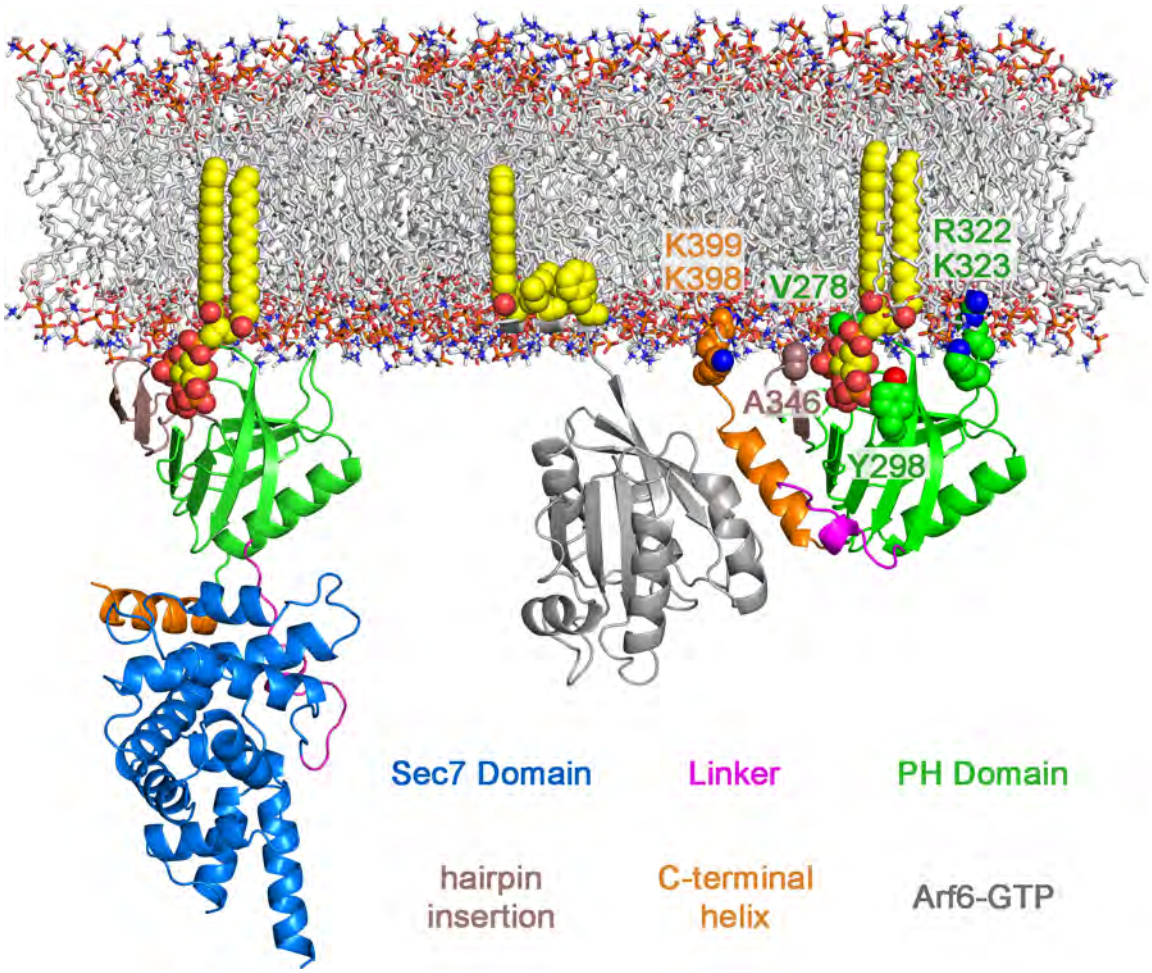

Figure II.13: Model for Integrated Relief of Autoinhibition and Membrane Recruitment of Cytohesins by Arf6-GTP. *Left*) Autoinhibited Grp1 structure (PDB ID 2R09). *Right*) Composite model for the active Grp1 complex with Arf6-GTP. The models are depicted in a common orientation relative to a model lipid bilayer based on the bound head group and known membrane targeting determinants. The N-terminal helix of Arf6 is modeled in an arbitrary orientation consistent with membrane partitioning. The myristoyl and diacyl glycerol moieties are modeled in configurations compatible with membrane insertion. The POPC bilayer membrane was derived from the coordinates of a molecular dynamics simulation [186].

membrane association of PH and histone domains exposes an allosteric site for Ras-GTP stimulation of GEF activity [158]. In BRAG2, association with membrane environments through nonspecific interaction at the PH domain stabilizes substrate binding and enhances catalytic activity [87]. In Cytohesins, membrane recruitment and allosteric relief of autoinhibition appear to be integrated through a unified structural mechanism whereby phosphoinositide binding primes Arf-GTP binding, which relieves autoinhibition by driving large scale conformational rearrangements that reposition the autoinhibitory elements to reinforce membrane partitioning and support Arf substrate activation by the Sec7 domain. Given the similarity with Sos, integration of autoregulatory and membrane targeting mechanisms in GEFs may be more prevalent than previously appreciated. Further investigation of the relationships between membrane targeting and allosteric activation will likely yield important insights into the molecular mechanisms underlying spatial-temporal-allosteric control of GTPase activation.

Materials and Methods

Constructs, Expression and Purification

Constructs were amplified using Vent polymerase, digested with BamHI-hf or NdeI and SalI-hf or XhoI, and ligated into pGEX-6P1 for expression as GST fusions or modified pET15b and pET21a encoding N- and C-terminal 6×His tags, respectively. Mutants were generated using the QuikChange II XL kit (Stratagene). Transformed BL21 (DE3) RIPL cells (Stratagene) were grown to an OD_{600} of 0.2-0.4 in 2×YT with 100 mg/L ampicillin and induced with 50 μM IPTG for 14-16 hours at 18-22°C. Cell

pellets were suspended in 50 mM Tris, pH 7.5 or 8.5, 150 mM NaCl, 2 mM $MgCl_2$, 0.1 mM PMSF, 0.2 mg/ml lysozyme, 0.01 mg/ml protease free DNAse I (Worthington), and 0.05% 2-mercaptoethanol. After sonication, lysates were supplemented with 0.5% Triton X-100 and centrifuged at 30,000×g for 1 hr. Proteins were purified over HisTrap Ni-NTA or glutathione-sepharose columns followed by ion exchange on HiTrap Q or SP and gel filtration on Superdex-75 (GE Healthcare).

For myr-Arfs, BL21 (DE3) RIPL cells (Stratagene) were co-transformed with pET21a encoding Arf6 and pBB131 encoding N-myristoyl transferase. Cells cultured in 2×YT at 37°C to an OD_{600} of 0.6-0.8 were supplemented with 50 μM sodium myristate and induced with 1 mM IPTG for 3 hrs at 37°C (Arf1/6-6×His) or 0.1 mM IPTG for 16 hrs at 22°C (untagged Arf1). After lysis in the absence of detergent and ultracentrifugation at 100,000×g, myr-Arf6 was extracted from the membrane fraction with buffer containing 1% Triton X-100. After ultracentrifugation at 100,000×g, the solubilized 6xHis protein was purified over Ni-NTA-Sepharose in buffer supplemented with 0.1% Triton X-100, concentrated, and stored in buffer containing 10 μM GDP and 0.1% Triton X-100.

Nucleotide Loading

Arf GTPases were incubated overnight at 4°C in 20 mM Tris, pH 8.0, 150 mM NaCl, 5 mM EDTA, and 1 mM DTT with a 10 fold molar excess of GTP, GppNHp or mantGDP. $MgCl_2$ was added to a final concentration of 10 mM and excess nucleotide

removed by gel filtration on Superdex-75 (GE Healthcare) or using a D-Salt column (Pierce ThermoFisher).

Determination of Nucleotide State

Nucleotide state for loaded and/or constitutively active Arf GTPases was determined by monitoring the change in intrinsic tryptophan fluorescence accompanying nucleotide exchange for GDP and GTP in the presence of 5 mM EDTA. The approach to equilibrium was monitored using a TECAN Safire microplate spectrometer, with emission and excitation wavelengths of 340 nm and 290 nm, respectively. The GTP-bound fraction was calculated as $\Delta I_{GTP} / (\Delta I_{GDP} + \Delta I_{GTP})$, where ΔI_{GTP} and ΔI_{GDP} were determined by a non-linear least squares fits to $\Delta I_{GTP} (1 - e^{-kobs\ t}) + I_{eq}$ and $\Delta I_{GTP}\ e^{-kobs\ t} + I_{eq}$, respectively. In cases where ΔI_{GDP} was too small to be measured, the GTP-bound fraction was conservatively estimated as >90%.

Liposome Preparation

Liposomes containing 67-70% palmitoyloleolylphosphatidylcholine (Avanti), 30% palmitoyloleolylphosphatidylserine (Avanti), and 3% PIP_2 or PIP_3 (Cell Signals) were prepared as follows. Lipids dissolved in 2:1 chloroform:methanol were mixed and dried under an argon stream. For GEF assays, lipid mixtures were rehydrated in 20 mM Tris HCl, pH 8.0, 150 mM NaCl, and 2 mM $MgCl_2$. Large Unilamelar Vesicles (LUVs) were prepared by extrusion through 0.1 μm filters (Avanti) after 10 freeze-thaw cycles in liquid nitrogen. For liposome partitioning, lipid mixtures were rehydrated in 20 mM

Hepes, pH 7.5, and 170 mM sucrose. Following 10 freeze-thaw cycles and extrusion as above, the liposomes were diluted 4:1 with 20 mM Hepes, pH 7.5, and 100 mM KCl, pelleted by ultracentrifugation at 100,000×g for 30 minutes, and resuspended at a total lipid concentration of 5 mM.

Nucleotide Exchange Kinetics

Nucleotide exchange was monitored as the fluorescence decrease upon mantGDP dissociation from Arf1NΔ17 in 20 mM Tris, pH 8.0, 150 mM NaCl, 2 mM $MgCl_2$ and 250 μM or 1 mM GppNHp. Grp1 with or without active Arf GTPases, liposomes, phosphoinositides, and/or head groups was formatted into 96 well half area microplates (Corning) and incubated for 16-24 hrs at 25°C. Reactions were initiated by addition of 1 μM Arf1NΔ17- or myrArf1-mantGDP and monitored using a Safire microplate spectrophotometer (Tecan) with excitation and emission wavelengths of 360 nm and 440 nm, respectively. Observed rates constants (k_{obs}) were obtained by fitting with $I_t = (I_0 - I_\infty) \exp(-k_{obs}\, t) + I_\infty$, where I_t, I_0, and I_∞ are the emission intensities at times t, t = 0, and as $t \rightarrow \infty$. Catalytic efficiencies (k_{cat}/K_M) were obtained from the slope of a linear least squares fit with $k_{obs} = (k_{cat}/K_M)[GEF] + k_{intr}$, where k_{intr} is the intrinsic dissociation rate constant. Half maximal activation constants ($K_{0.5}$) were determined by fitting with the hyperbolic function $k_{obs} = (k_\infty - k_0)\,[Arf6\text{-}GTP] / (K_{0.5} + [Arf6\text{-}GTP]) + k_0$, where k_0 and k_∞ are the values of k_{obs} at t = 0 and as $[Arf6\text{-}GTP] \rightarrow \infty$. For experiments in which the effective $K_{0.5}$ was comparable or less than the [GEF], the observed rate constants were fit with a quadratic function $k_{obs} = (k_\infty - k_0)\,\{b - (b^2 - 4\,[Arf6\text{-}GTP] / (n\,[GEF]))^{1/2}\} + k_0$,

where b = 1 + [Arf6-GTP] / (n [GEF]) + $K_{0.5}$ / (n [GEF]) and the binding stoichiometry n was fixed at 1.

Surface Plasmon Resonance

CM5 sensor chips were docked with Biacore S3000 or T100 instruments (GE Healthcare), activated, and coupled with anti-GST according to the manufacturers instructions. Proteins were dialyzed into 20 mM Tris, pH 8.0, 150 mM NaCl, 2 mM $MgCl_2$, and 0.005% P-20 surfactant. Equivalent molar quantities of GST-Grp1 constructs or GST were loaded on the sample and reference channels, respectively. 6×His-tagged Arf6-GppNHp or Arf6NΔ13 Q67L was injected for 60 s at 10 µl/min. After alignment, baseline correction, and reference subtraction, equilibrium responses (R_{eq}), which correspond to the equilibrium signal level, were determined from the average response over a 30-50 s range at the end of the injection. Dissociation constants (K_D) were determined by fitting with $R_{eq} = R_{max}$ [Arf6] / (K_D + [Arf6]), where R_{max} is the value of R_{eq} as [Arf6]→∞.

Crystallization and Structure Determination

Ternary complexes were formed by incubating 6×His $Grp1_{247-399}$, 6×His Arf6NΔ13 Q67L, and IP_4 in a 1:1:1.2 molar ratio at a total concentration of 10 mg/ml for 16-18 hrs at 25°C. Crystals were grown by vapor diffusion from 50 mM Tris, pH 8.8, 16-18% PEG 4000, and 0.2 M sodium citrate, transferred to a cryoprotectant solution (50 mM Tris, pH 8.8, 20% PEG 4000, 10% glycerol, and 0.2 M sodium citrate), and flash

frozen in liquid nitrogen. The crystals are in the space group $P4_32_12_1$ with cell dimensions of a = b = 56.6 Å, c = 274.4 Å and contain a single complex in the asymmetric unit. Diffraction data were collected at the Brookhaven NSLS X25 beamline and processed/scaled using HKL2000 [187]. The structure was solved by molecular replacement with Phaser using the Grp1 PH domain (1FGZ) and Arf6-GTPγS (2J5X) structures as search models. A σA-weighted $2wF_o$-DF_c map calculated after interleaved atom-updating and refinement with ARP/wWARP [188] and Refmac-5 [189] was used for automated model building with Buccaneer [190]. IP_4, GTP and solvent molecules were added using Coot [191]. The structure was refined by iterative rebuilding with Coot, atom updating with ARP/wARP, positional refinement with Refmac5, and simulated annealing with Phenix [192]. Structural figures were generated using PyMOL.

Statistical Analysis

Values and errors represents mean ± standard deviation for 2-4 independent measurements.

Chapter III: Development of Data Processing and Analysis Techniques for Size Exclusion Chromatography Small Angle X-ray Scattering

Abstract

Small Angle X-ray Scattering (SAXS) is a widely-used technique for structural analysis of macromolecular systems in solution. Recently, size exclusion chromatography in line with SAXS (SEC-SAXS) at third generation beamlines has helped bypass some of the traditional difficulties in data collection with challenging samples, but data analysis procedures still rely heavily on somewhat subjective trial and error methods unsuited for large data sets produced by SEC-SAXS. Presented here are novel methods for data processing and analysis incorporated into a Python-extensible graphical software package. Intensity normalization, buffer correction, Guinier analysis, and constant subtraction procedures have been optimized using data for the GTPase Arf6 and its Guanine Nucleotide Exchange Factor (GEF) Grp1.

A method of scattering profile reconstruction was developed using singular value decomposition of the sample data and linear combination of basis components representing the space of sample and buffer scattering to process data as an alternative to direct buffer subtraction. This technique was especially useful when characterizing partially overlapping oligomeric species or in cases where suitable buffer regions are not available. Additionally, constant background subtraction from SAXS data was systematically explored so as to accurately generate *ab initio* shape envelopes for known samples. Overall, these studies highlight the strengths of SEC-SAXS as a tool for structural analysis, and unify techniques for data processing and analysis between this and other data collection methods.

Introduction:

X-ray crystallography has remained an unparalleled technique in structural biology for its ability to visualize proteins at the atomic level. Indeed, these static structures often accurately inform about mechanisms of action at the whole protein, domain, and atomic levels. However, because this technique requires crystallizing the molecule in a non-physiological environment or context that often includes high concentrations of salts, truncation of native proteins, and/or incorporation of bulky external polypeptide tags, it can be an imperfect, incomplete representation of the macromolecular structure, and provides limited information on the conformational flexibility in the system. Alternative techniques such as nuclear magnetic resonance (NMR) have size limitations that exclude many large, flexible systems which are often highly sought after structural targets. For these reasons, Small Angle X-ray Scattering (SAXS) has emerged as a useful tool to obtain native-like structural information, albeit at lower resolution, on a wide variety of biological systems [193-196].

In a typical SAXS experiment, buffer and protein samples are exposed to synchrotron radiation under flow and the measured intensities at various scattering angles (θ) are radially averaged to create one dimensional intensity profiles as a function of the momentum transfer vector q (defined herein as $|q| = 4\pi \sin\theta / \lambda$, where 2θ is the scattering angle and λ is the wavelength in Å). Although dominated by solvent, the scattering profile of the sample will contain low resolution information about the protein, which is an ensemble of randomly distributed orientations in solution. Extraction of the protein contribution to scattering can be obtained through direct subtraction of the buffer

trace from the protein trace after normalization by the incident (I0) or transmitted (I1) beam intensity, and may be aided by water calibration using Wide Angle X-ray Scattering (WAXS) [197-200] or by inter-protein correction [201]. These corrected scattering curves can then be used for subsequent analysis.

Because SAXS measures bulk scattering in solution rather than specific diffraction from highly pure and static crystalline samples, sample quality is the primary factor affecting the reliability of the results. Scattering intensity (I) is directly proportional to the molecular weight of the sample and is somewhat difficult to parse from total scattering signal of other components. In practice, this manifests as large contributions from any contaminants due to aggregation or degradation and high sensitivity to buffer matching procedures. Buffer matching, reduction of concentration-dependent effects, and using homogeneous, well-folded samples are necessities of any experiment determining protein structure by SAXS [193, 196]. Due to diversity in specific requirements for biological macromolecules, however, there is little standardization in the way SAXS experiments are set up with regards to these three criteria, and often limitations of particular systems dictate experimental procedure [198, 199, 202]. Recently, Size Exclusion Chromatography in line with SAXS (SEC-SAXS), has been integrated at several synchrotron beamlines to ensure highly pure, monodisperse samples [203-206]. The clear advantages of this method have made this technique ideal for studying many challenging macromolecular systems.

Regardless of the collection strategy, sample quality is often hard to measure, and the presence of protein impurities or other problematic artifacts is not always evident in

the raw data, requiring detailed analysis of all experiments. Graphical analysis techniques such as Guinier, Kratky, and Porod plots provide a measure of sample quality and more quantitative information about overall size and shape [193]. Current analysis tools, while functional, are time-consuming and rely on much manual manipulation and interpretation of the data, restricting the amount of data processing that can be performed during limited synchrotron beam time, and potentially producing ambiguous results. SEC-SAXS further complicates this due to the large quantity of data produced, and the fact that buffer and protein samples have long time gaps between their exposure. There is a need for flexible, graphically based tools incorporating objective methods for analysis of data quality and optimization of buffer subtraction.

Beyond graphical analysis, low resolution (typically 5-15 Å) *ab initio* shape envelopes can be generated using spherical harmonic approximation of shapes based on the overall pairwise distance distribution function P(r) [207-212]. The Svergun group has developed a large suite of programs for calculation of P(r) distributions and generation of models [213]. DAMMIF and GASBOR use *ab initio* dummy atom or dummy residue models as an over-parameterized approach to fit the data using simulated annealing [214-216]. These analyses are generalizable to a wide array of molecules, and give useful insight to an overall structure for further study, but due to the large number of parameters they are potentially ambiguous and require many trials and averaging for confidence in the result [217]. Moreover, while many of the tools developed for shape reconstruction using SAXS analysis were developed using highly idealized or theoretical systems and optimized using experimental data from traditional methods, little work has been done to

address how best to use them with more modern SAXS data collection techniques and non-ideal data sets, often having high confidence, but more irregular molecular architectures.

This chapter describes widely applicable, novel approaches for SAXS data processing and analysis developed using SEC-SAXS. A software package called Data Evaluation and Likelihood Analysis (DELA) has been developed and integrated with python-based scripts to intuitively perform automated SAXS analysis such as data manipulation and correction as well as visualization and optimization of Guinier, Kratky, and Porod plots [218]. Using this as well as other standard programs we have performed intricate analysis troubleshooting beam intensity correction, buffer correction, constant subtraction, P(r) calculation, and dummy atom shape envelope construction using data from samples with known structures from the Arf GTPase Arf6 and its Guanine Nucleotide Exchange Factor (GEF) Grp1.

For beam intensity normalization, a scouting procedure was developed to determine the best method based on SEC-SAXS total intensity chromatograms. Whereas most SAXS experiments include normalization by transmitted intensity at the beam stop, the best correction occurred from normalizing by the incident and transmitted intensities to account for beam intensity fluctuations obscured in the transmitted recording. Next, a novel buffer correction method was developed that uses singular value decomposition and linear combination of protein and buffer scattering components. This reconstruction performs markedly better compared to simple buffer subtraction, and importantly allows the user to determine the homogeneity of samples and efficiently process large SAXS

data sets from any type of data collection with ease. Multiple, low-abundance species could also be reliably analyzed in a single experiment. Finally, we demonstrate an optimized strategy for overall constant subtraction and the use of current dummy atom modeling programs that retains high q SAXS data and provides reliable shape envelopes for known and novel samples. Overall, our analysis strategies advance the capability of SAXS to confidently analyze a variety of potentially low-abundance macromolecules through real-time purification and vastly redundant data collection supported by sophisticated analysis techniques capable of using this wealth of information.

Results

SEC-SAXS Data Collection, Processing and Analysis

SEC-SAXS provides ideal conditions for sample purity and data quality. In the SEC-SAXS setup at the Advanced Photon Source (APS) at Argonne National Labs in Argonne, IL, a commercially available Fast Performance Liquid Chromatography (FPLC) gel filtration column is set up for flow in line with SAXS optics. A Superdex 200 10/30 (24 mL) or Superdex 10/300 mm (3 mL) column were used for these experiments. When run at 0.5 ml/min at room temperature, the 3 mL column offers significantly higher throughput (~10 min per sample), with negligible decreases in resolving power. Scattering from buffer and protein eluting from the column was sampled with one second exposures every five seconds. After radial averaging, the result is a series of 100-200 SAXS curves corresponding to buffer and protein samples (Figure III.1).

As a consequence of the amount of data produced by SEC-SAXS, data processing becomes a bottleneck. Thorough assessment of data quality requires analyzing overall raw data, mean scattering, buffer subtracted data, and Guinier and/or Kratky plots. Even with the high throughput of SEC-SAXS data collection, current analysis tools present challenges for checking every sample during allotted beam time. To expedite this process, we developed a Python-extensible, native Mac OSX application called Data Evaluation and Likelihood Analysis (DELA), which allows users to quickly and objectively process and analyze data from a variety of experimental methods including SAXS using Objective-C coded source functionality extended through an embedded Python interpreter, with a suite of modules bridging user-modifiable Python scripts with the main application. The application bundle after extension with SAXS-specific scripts supports importing/exporting, processing, and analysis of large quantities of raw scattering curves of any size (limited only by available computer memory) in an intuitive, graphical interface. Processed data and results can be archived as a document, exported for further analysis, and publication quality plots saved in common image formats.

To aid in development of SEC-SAXS data processing and analysis, we used SAXS data for an Arf GTPase Arf6 and one of its Guanine Nucleotide Exchange Factors (GEF), Grp1. The Arf6 construct (residues 14-175) consists of a core G-domain well documented in structural studies (PDB ID 2J5X). A Grp1 construct (residues 63-399, $Grp1_{63\text{-}399}$) composed of two uniquely shaped domains that adopt a static, elongated structure seen in X-ray crystallography (PDB 2R09) as well as a dimeric construct

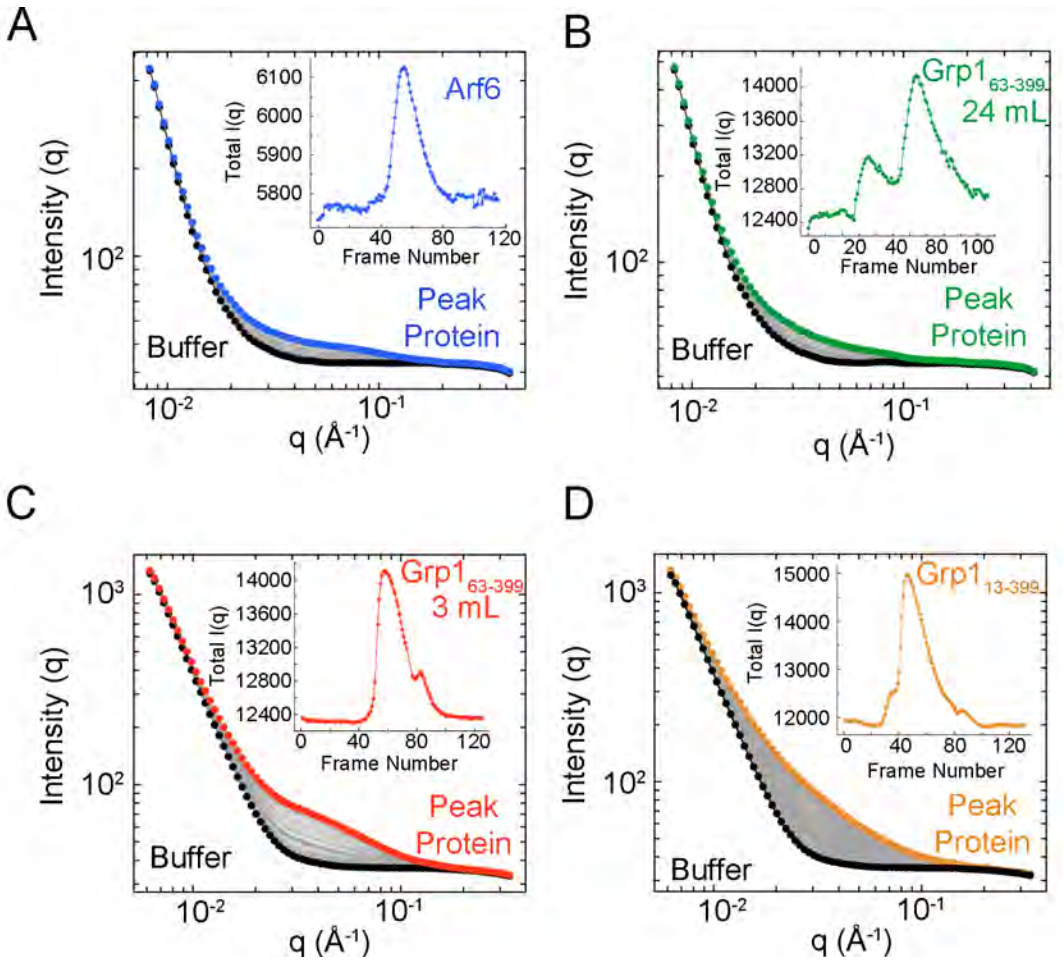

Figure III.1: SEC-SAXS Analysis for Arf6 and Grp1. Raw Intensity profiles and total intensity chromatograms (inset) for Arf6 (A), $Grp1_{63-399}$ 24mL (B), $Grp1_{63-399}$ 3mL (C), and $Grp1_{13-399}$ (D).

including a heptad repeat region (residues 13-399) were analyzed. All constructs have an N-terminal His_6 tag for these studies.

Figure III.1 shows the analysis of SEC-SAXS data for the three constructs. Arf6 was analyzed using a 24 mL column, $Grp1_{13-399}$ using a 3 mL column, and $Grp1_{63-399}$ using both columns ($Grp1_{63-399}$ 3mL, $Grp1_{63-399}$ 24mL). Scattering intensities were integrated for each data set to generate a total intensity chromatogram as a function of the data frame during the experiment. One or more regions of the chromatogram representing data sets of interest can be selected and the raw data representing them plotted using a graphical selection tool and python scripts (Figure III.1). Due to larger loading concentrations on the 24 mL column, a concentration-dependent Grp1 dimer was isolated on the 24 mL column, but not the 3 mL column. In this way, high and low signal experiments analyzing monomeric and dimeric species of the same sample could be analyzed. Overall, monomeric protein of each construct was separated from either dimers or contaminants accumulating during sample preparation that would not have been separated in traditional SAXS experimental procedures.

Normalization using incident/transmitted beam intensities

The profiles in Figure III.1 are displayed after correction by incident (I0) and/or transmitted (I1) beam intensities. To determine the best normalization procedure, profiles of I0, I1, and total sample intensity for each sample were compared, as shown in the example with $Grp1_{63-399}$ 3mL (Figure III.2A and B). All three measurements showed fluctuations in intensity throughout the experiment. I1 and the total intensity both

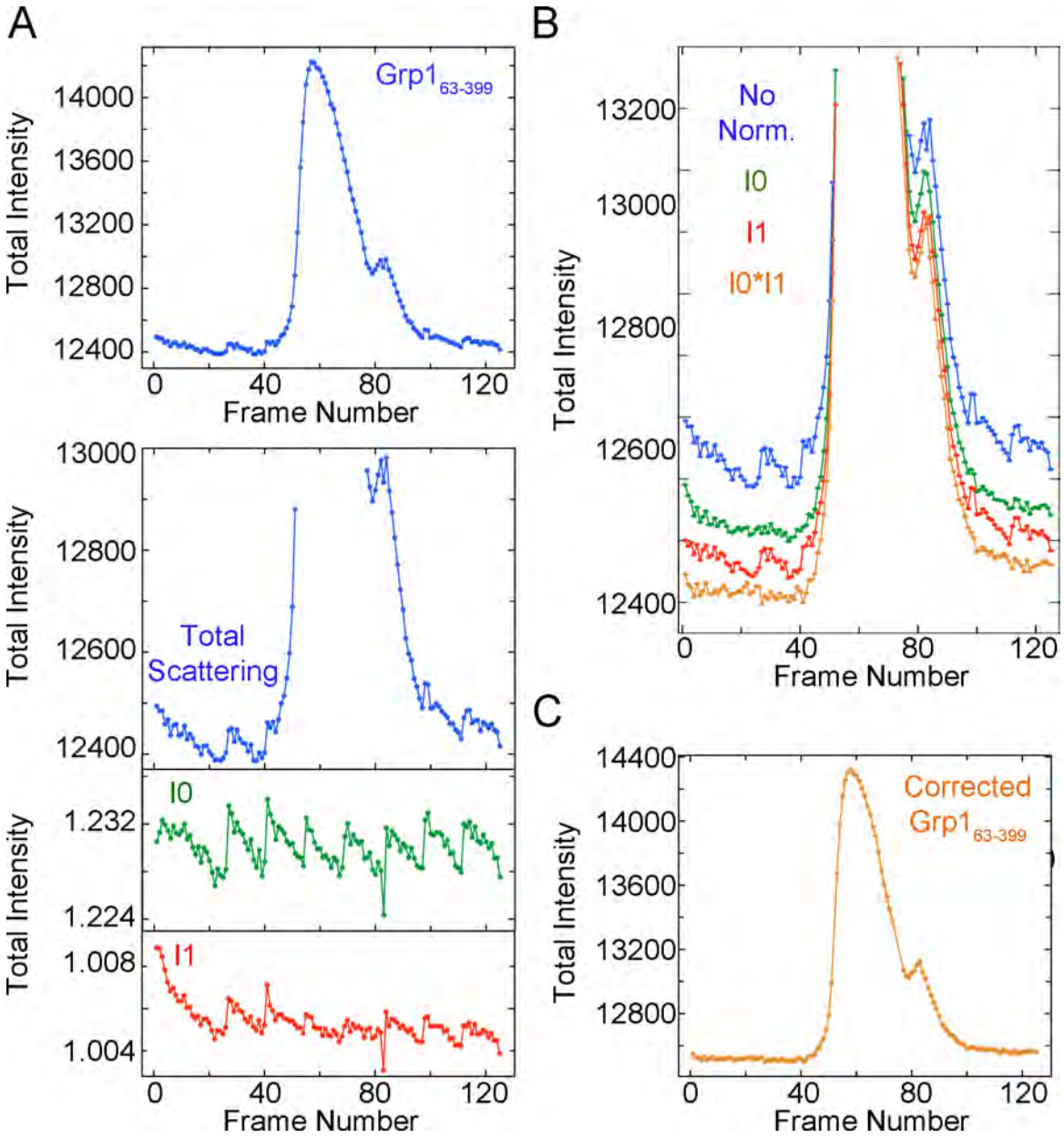

Figure III.2: Data Normalization of $Grp1_{63-399}$. A) Total intensity chromatogram of the $Grp1_{63-399}$ sample (top) and zoomed in view of sample, I0, and I1 intensities (bottom). B) Baseline intensities after indicated normalizations, offset for clarity. C) Final total intensity profile after normalization

showed an initial downward slope in the total intensity, while I0 lacked this feature but had more pronounced variations reflecting periodic beam injection and decay. Normalization corrections were scouted by correcting the sample data with I0 or I1 in varying weights through a scriptable, iterative process. Normalization by I0 alone reduced the baseline scattering fluctuations, and normalizing by I1 alone reduced the baseline slope somewhat.

Where applicable, normalization procedures were empirically determined to yield a normalized total scattering chromatogram with the least contribution from beam intensity artifacts. Typically, this was either a weighted, scaled average of I0 and I1 or a normalization by both factors (Figure III.2B). From a practical standpoint, normalization by the product of I0 and I1 was found to yield consistently better results than normalization by either I0 or I1 alone. Importantly, the details of the correction process did not change the overall scattering relationships, and samples corrected multiple ways had similar results upon further analysis, but rather produced clean total intensity chromatograms to analyze. This normalization process clearly identifies and corrects for significant components in both the incident and transmitted beam intensities, which may be especially important for low signal-to-noise/background scattering data.

Buffer Subtraction Scouting

To investigate data quality, initial scouting of buffer regions for subtraction was performed, and data analyzed by a Guinier plot. The Guinier plot (log I(q) vs. q^2) provides useful structural information at low q angles. For the q range satisfying the

Guinier approximation, this plot has a linear dependence with a slope dependent on the overall size and shape of the molecule [193, 219] and proportional to the radius of gyration, R_g. Since the Guinier plot is independent of concentration, it can be used as a measure of data quality across concentrations and provides information on sample multimerization, aggregation, and what has commonly been termed ‘interparticle interference’, thought of as the exclusion of solvent at high protein concentrations through interaction or repulsion of protein molecules.

Bulk increases in concentration-normalized scattering on a Guinier plot are indicative of changes in oligomeric state upon concentration. Alternatively, samples with high amounts of aggregation will produce upward trends near I(0), and those with interparticle interference will slope downwards in a similar range [193]. Notably, the linearity of the Guinier region is sensitive to buffer subtraction artifacts, and even slight mismatches (< 1%) will produce positive or negative features that are difficult to distinguish from those associated with small amounts of a high molecular weight aggregates or so-called interparticle interference in traditional SAXS experiments. With SEC-SAXS, high molecular weight aggregates can be effectively eliminated, while, buffer matching and interparticle interference remain important considerations.

In SAXS experiments outside of SEC, buffer subtraction would include independently averaging beam-normalized buffer and protein scattering curves, after rejection of outliers, then performing direct subtraction of the averaged buffer protein curves, with or without concentration-dependent buffer scaling to account for the fraction of the sample solution volume displaced by the protein. Similarly, an initial buffer

subtraction strategy for the $Grp1_{63-399}$ 3mL and Arf6 data sets consisted of averaging and subtracting three experimental buffer regions from protein samples comprising the peak of the monomer elution (Figure III.3A and C). This is expedited by the ability within DELA to directly plot representative scattering curves from total intensity plots and perform sample averaging, buffer subtraction, and an automated Guinier analysis.

Unexpectedly, despite the high quality of the data and low protein concentration, direct subtraction of buffer produced a sharp downward curve in the low q region (Figure III.3B and D). Omission of aberrant data points from the fit produced Guinier fits with R_g values of 25.8-27.6 Å for $Grp1_{63-399}$ and 16.4-16.6 Å for Arf6, similar to theoretical values from the crystal structures produced by CRYSOL (28.0 Å and 16.7 Å, respectively). Although some of these artifacts could be partially ameliorated by choosing specific regions for analysis as with Arf6 (Figure III.3D), this trial and error method of data analysis was less than ideal, as it was arbitrary and did not take full advantage of the large redundancy of data in SEC-SAXS. More generally, it does not account for incompatibility of various buffer and protein regions due to either large time lapses or slight compositional differences between buffer and protein sample analysis, imperfect beam-normalization, and/or other sources of experimental error, all of which can be reduced but not completely eliminated, particularly in the case of dilute or weakly scattering samples. These are inherent issues in SEC-SAXS, but also problematic for traditional SAXS data collection, where data for protein and buffer samples are collected at different times with perhaps less than ideal buffer matching.

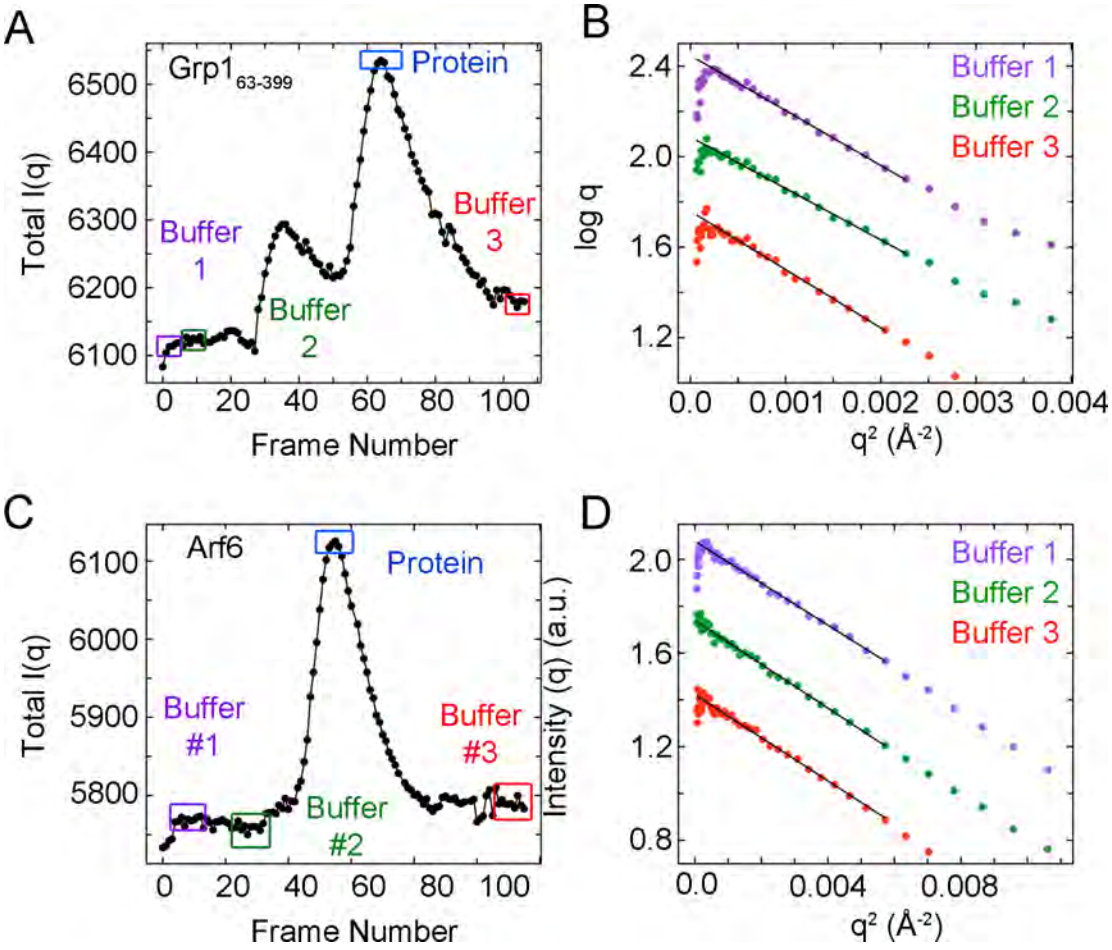

Figure III.3: Buffer Subtraction Trials for Arf6 and $Grp1_{63-399}$. Intensity chromatograms for $Grp1_{63-399}$ 24mL (A) and Arf6 (C) indicating regions of buffer and protein used in buffer subtraction. Buffer subtraction results are shown for each buffer region after sample averaging for $Grp1_{63-399}$ 24mL and Arf6 in B and D, respectively.

Assessment of Homogeneity by Singular Value Decomposition

To assess overall information content of samples for buffer correction, Singular Value Decomposition (SVD) was calculated on beam-normalized scattering curves as in Figure III.2 from part (e.g. peak of interest and buffer regions) or in some cases all of the SAXS chromatogram. SVD has been used in protein folding and SAXS model fitting in cases where a measure of the uniqueness of a scattering result is desired [220, 221]. In brief, SVD is a matrix algebra method that is particularly useful for determining the minimum components required to accurately represent data with a high degree of redundancy, such as protein peaks and buffer regions in SEC-SAXS, which differ primarily by the concentration-dependent scattering contribution of the protein. As applied to SEC-SAXS, the data are represented as an M • N matrix with M columns, each corresponding to a different scattering curve consisting of N intensity values over the observed q range, and deconstructed into orthonormal basis vectors (columns of an M • N matrix U), singular values (elements of an N • N diagonal matrix S), and orthonormal coefficients (columns of a transposed matrix V) that, after weighting by the singular values, specify the linear combination of basis components required to exactly reconstruct each scattering curve in the original matrix:

$$A(M\ x\ N) = U(M\ x\ N) * S(N\ x\ N) * V^T(N\ x\ N)$$

A particularly useful property of singular value decomposition is the rotation of the matrices so as to successively maximize the contribution of each basis vector, creating a rank ordered series of basis vectors that maximizes the contribution of the first column of U (i.e. it is the most representative) as reflected in its corresponding singular value, then

maximizes the contribution of the second column to the remainder (also reflected in its corresponding singular value), and so on. Thus, the singular values are indicative of the relative contribution or weight of each column of U to the entire set of data in S, with the most significant (i.e. highly weighted) U columns having reduced contributions from random noise in the data. The reduction in noise is immediately apparent when considering that the first column of U is simply a normalized summation of the individual scattering curves.

The use of SVD for data analysis was tested with the $Grp1_{63\text{-}399}$ 24mL data set (Figure III.4). Using a selection of regions of highly pure protein and solvent chosen from the plot of total scattering intensities, the data could be represented with two dominant components, as evidenced by the magnitude of the singular values and by the autocorrelation of the columns of U and V. This analysis was similar when using data from the whole protein peak, and inclusion of dimer or contaminant portions of the experimental data produced an additional component seen in the autocorrelation plots (Figure III.4C). Using this technique combined with SEC-SAXS, one can objectively assess the quality of data in a specific range and across large amounts of data without the need to scale for protein concentration or compare multiple rounds of buffer subtraction and graphical analysis.

Indirect Buffer Subtraction by Linear Combination of SVD Basis Vectors

The preceding analysis suggests that regions of data consisting primarily of buffer and a single homogenous protein species can be identified using SVD, even in cases

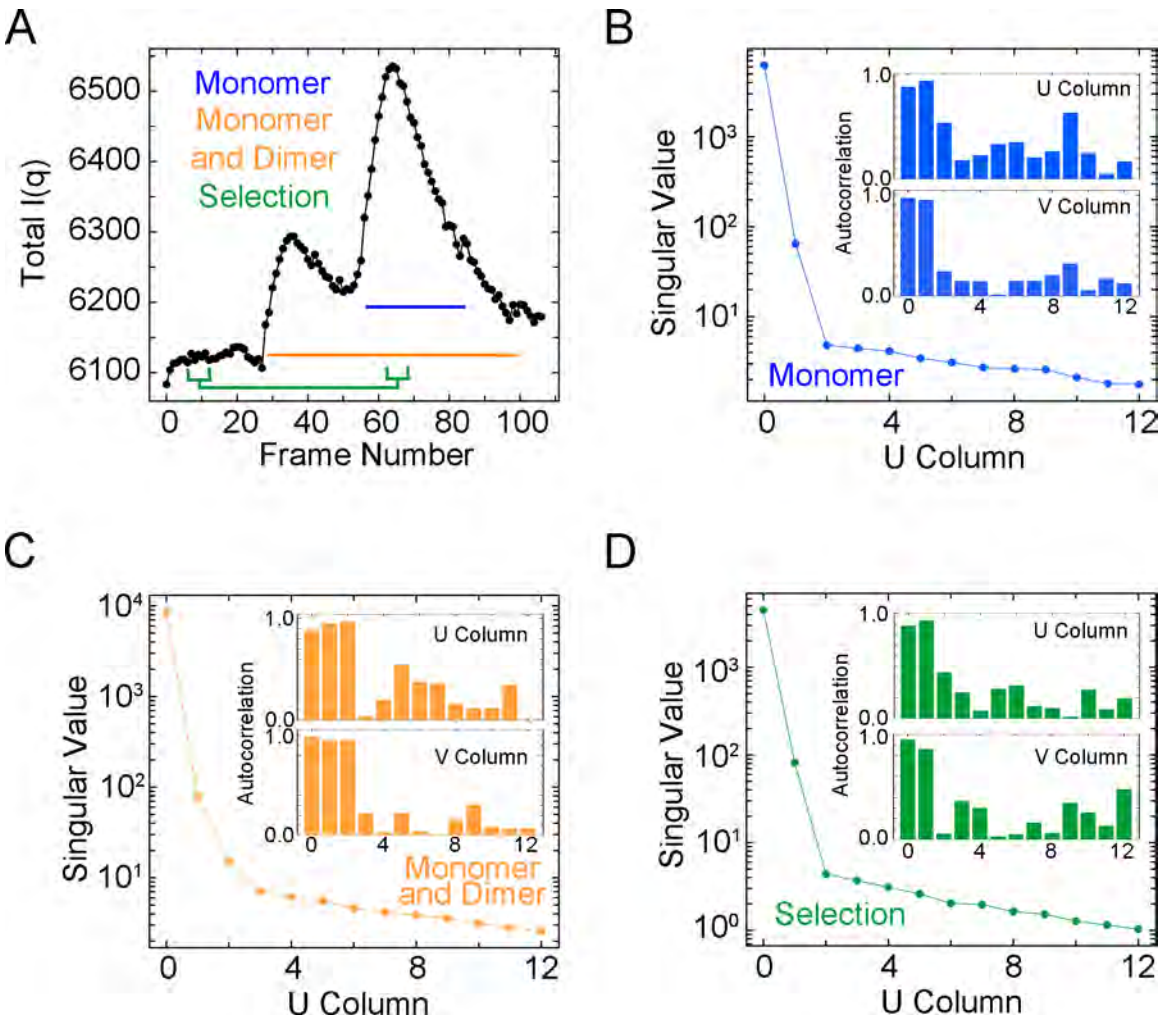

Figure III.4: Singular Value Decomposition (SVD) of $Grp1_{63\text{-}399}$. A) Schematic indicating regions from $Grp1_{63\text{-}399}$ 24mL used in SVD analysis. B-D) Singular Values and Autocorrelations (insets) of Monomer (B), Monomer and Dimer (C), and Selected protein and buffer regions (D).

where a suitable buffer region is lacking. The finding that, apart from noise, the scattering data from part or all of an eluting peak with or without additional buffer regions can be accurately represented by two components implies that the two dimensional space occupied by the pure protein and buffer components can be adequately represented by the two significant basis vectors of U. It then follows that the protein component should in principle correspond to a linear combination of U0 and U1, with weights c0 and c1:

$$I_q(protein) \propto c0 \times U0 + c1 \times U1$$

If an arbitrary scale is assumed, one of the coefficients (e.g. c1) can be set to one, and the problem reduces to finding the best value for the unknown coefficient (c0). An approximate value for c0 can be empirically determined by visual inspection of the Guinier region, which is expected to be linear for the optimal coefficients but deviate positively or negatively otherwise. Deviations from linearity are readily detected using weighted or unweighted merit statistics such as R^2, suggesting an optimization procedure that uses linearity of the Guinier region as a metric for determining the best value for c0.

An example for the $Grp1_{63\text{-}399}$ 3 mL sample is shown in Figure III.5. The two highest signal components from SVD were combined using iterations of constants within a specified range, and the resulting curves subjected to a Guinier analysis using a Python script for determination of the optimal value of c0 based on R^2 has been integrated into DELA. The best fitting combination can then be combined. Non-uniform errors were estimated by Savitzky-Golay smoothing using a 3rd order polynomial with a window size of 11 points [222]. Reconstruction of both Arf6 and monomeric $Grp1_{63\text{-}399}$ data through

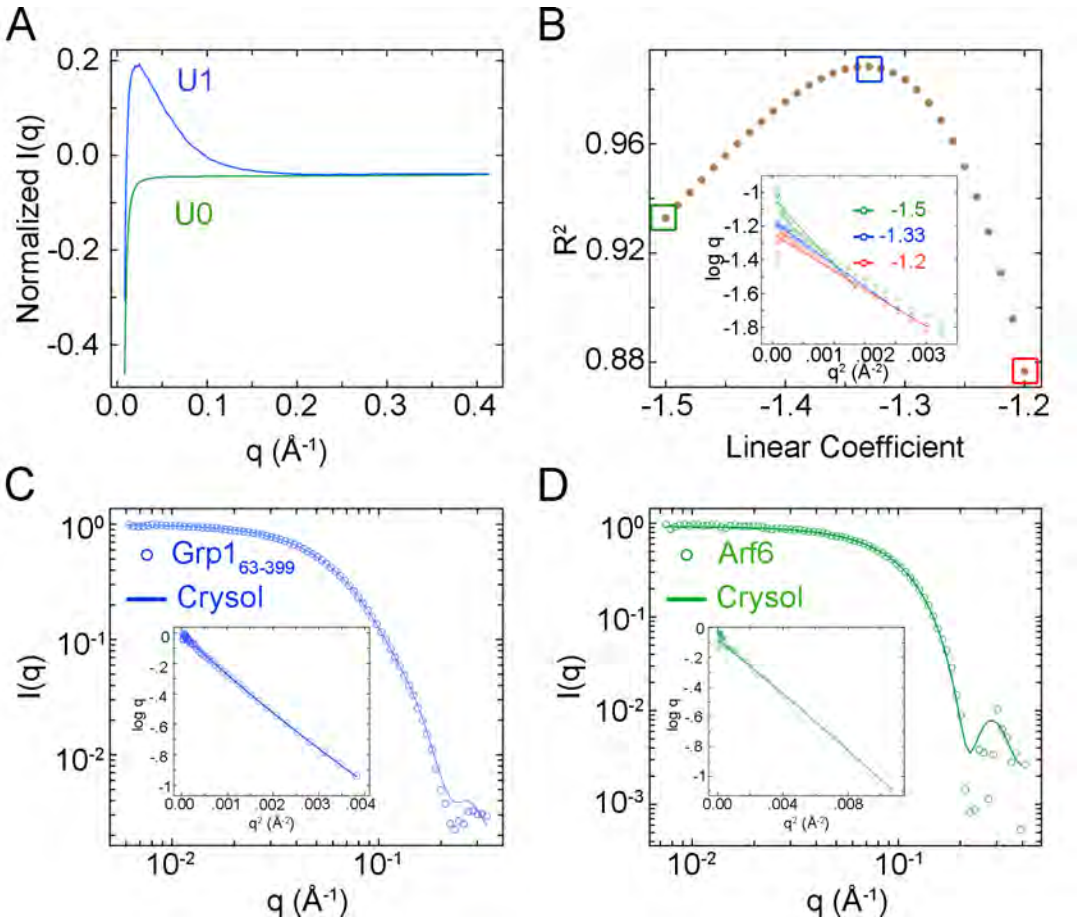

Figure III.5: Linear Combination of SVD Components. A) Scattering profiles of two most abundant components from SVD of the $Grp1_{63-399}$ 3mL sample. B) Linear coefficient analysis from components in A and example Guinier plots and fits from specific coefficients (inset) as indicated by colored boxes. C and D) Corrected scattering curve and Guinier plot (inset) after optimal linear combination for $Grp1_{63-399}$ 3mL (C) and Arf6 (D). Solid line in both depicts scattering calculated from the crystal structures of $Grp1_{63-399}$ (PDB ID 2R09) or Arf6 (PDB ID 2JFX) using CRYSOL with buffer subtraction.

this method, termed SVD-LC, produced a high signal to noise reconstruction of the putative protein scattering that was nearly identical to the theoretical scattering for the crystal structure of the same construct calculated with CRYSOL (Figure III.5C and D).

Indirect buffer subtraction by SVD-LC was also used to analyze the low abundance dimer seen in the $Grp1_{63-399}$ 24mL experiment. Gel filtration experiments with a $Grp1_{63-399}$ I307E mutant that showed increased dimer formation confirmed that the dimer was concentration dependent, seen maximally as about 10% of the overall sample in the presence of reducing agent. An additional oligomeric peak representing less than 1% of the sample was observed with loading concentrations above 200 μM. Traditional buffer subtraction with these low signal samples produces noisy corrected curves as described, but the SVD-LC approach allowed unambiguous reconstruction of the protein contribution with a sample concentration estimated as < 1 mg/ml at the peak. The resulting reconstruction closely matched theoretical scattering for the crystallographic dimer for this construct (Figure III.6).

Constant Background Subtraction Optimization for Ab Initio Shape Determination

As illustrated by the examples described in the previous section, SVD-LC is capable of producing high quality reconstructions of the protein scattering that are optimized with respect to linearity of the Guinier region. The quality of the Guinier region is important for determination of the overall size and molecular weight of macromolecules by SAXS. However, the linearity of the Guinier region is relatively insensitive to the data at high q and thus the optimization based solely on the Guinier

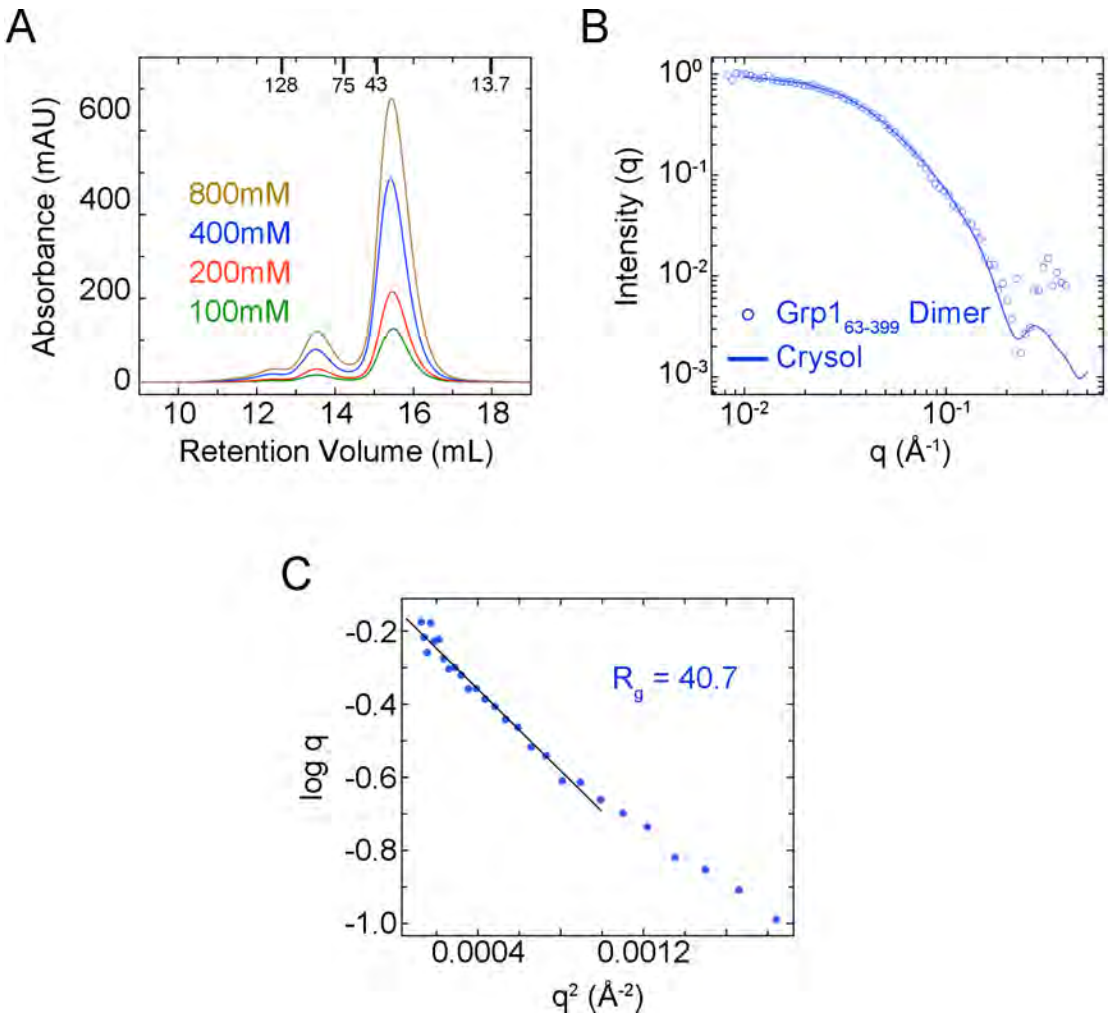

Figure III.6: Characterization of the Grp1$_{63-399}$ dimer species by SEC-SAXS. A) Gel Filtration analysis of Grp1$_{63-399}$ I307E, with protein concentrations at loading as indicated. B) Corrected scattering curve of dimer region from SEC-SAXS analysis of Grp1$_{63-399}$ 24mL. Solid line is the theoretical scattering calculated from the crystallographic dimer of Grp1$_{63-399}$ from CRYSOL with buffer subtraction. C) Guinier plot and fit from data in B.

region may incur small residual deviations that become relatively more significant with increasing q. These small systematic errors are analogous to the problem of buffer matching for direct subtraction and may share some common underlying causes (e.g. imperfect beam-normalization). As with direct buffer subtraction, small deviations in final curves reconstructed by SVD-LC on the order of a percent or less are difficult to diagnose and eliminate. Nevertheless, the effects of such errors can be partially mitigated by constant background subtraction, as implemented by default in many programs for *ab initio* or rigid body modeling. While many programs in the ATSAS package automatically determine a constant for subtraction, in some cases, such as ab initio modeling by DAMMIN/DAMMIF, automatic determination of a constant for background subtraction is not performed if sufficiently high q data are not available.

To account for this, the scattering intensity decay dependent on the overall volume and shape of the molecule were used [223]. Over a portion of the scattering data, referred to as the "power law regime", purely spherical molecules have scattering decay proportional to q^4 (Porod's law) [224], with more complex topologies following lower power relationships such as $q^{5/3}$ for random coils. Unfolded proteins show little to no Porod scattering [193, 196]. These relationships underscore the importance of accounting for solvent scattering in this region for an assessment of volume and thus structure determination. One approach is to collect wide angle X-ray scattering (WAXS) and SAXS data, using the WAXS data as a guide to determine a scale factor to reduce the solvent contribution, seen as a ring of intensity at $q \sim 2.0$ Å^{-1} [122, 200]. In effect, this is equivalent to a constant subtraction scaled to an independent scattering value measured

for each sample. In the absence of WAXS data, the presence of 'contamination' from high q solvent intensity can be seen graphically in Kratky plots (q^2*I(q) vs. q) and Porod plots (q^4*I(q) vs. q), which for well folded proteins, produce a peak related to the overall volume of the protein, referred to as the Porod Volume defined by 1st and 2nd Porod minima [225]. An accurate Porod volume is necessary for accurate calculation of the pairwise distribution function and therefore *ab initio* shape construction. Initial analyses performed on a subset of structures in order to develop algorithms for this calculation found data up to $q = 8/R_g$ was adequate to reconstruct the first and second Porod volume [217]. However, as many proteins follow lower power scattering decay than the q^4 Porod scattering relationship, power law regimes may be more or less extended compared with simple models, depending on the actual shape of the molecules investigated.

To explore the effects of constant background subtraction, protein scattering curves for $Grp1_{63-399}$ 3mL and $Grp1_{13-399}$ were reconstructed with SVD-LC, normalized to 1.0, and plotted on a Porod plot (Figure III.7). Although the data displayed peaks indicative of well folded protein, samples lacked a clear 2nd Porod Minimum. Constant subtraction was performed iteratively using constants from 0.001-0.008, with significantly improved definition of the 2nd Porod minimum at with constants between 0.004-0.005. Oversubtraction produced data with negative intensities and severely warped Porod peak shapes. For $Grp1_{63-399}$ 3mL, the 2nd Porod minimum was at about 0.22 $Å^{-1}$, well within the $8/R_g$ cutoff of 0.289 $Å^{-1}$ for this construct (R_g = 27.7, Figure III.7A). In contrast, the dimeric $Grp1_{13-399}$ had an R_g of 50.4 Å, but a 2nd Porod minimum at ~ 0.25 $Å^{-1}$, well outside the $8/R_g$ cutoff of 0.159 (Figure III.7B). Overall, these data

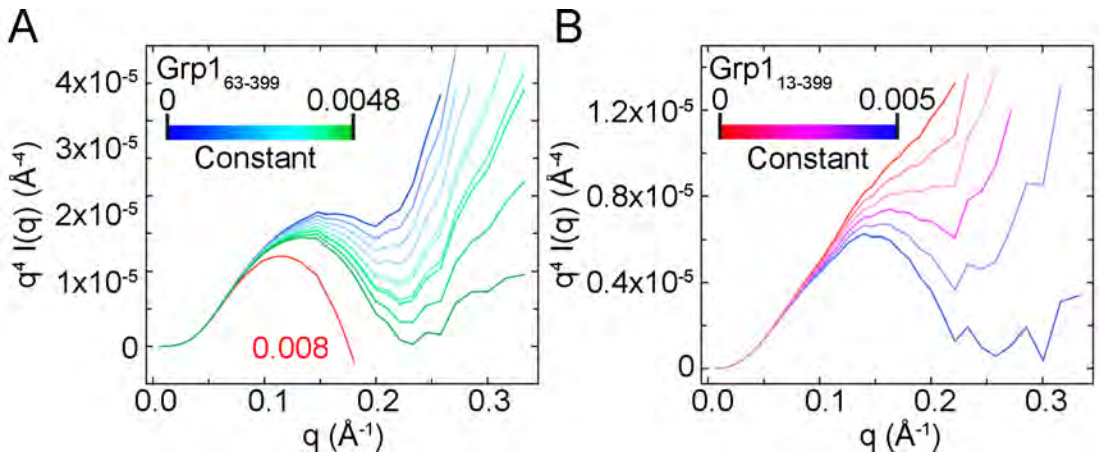

Figure III.7: Constant Subtraction Scouting by Porod Plot. Porod plots of $Grp1_{63-399}$ (A) and $Grp1_{13-399}$ (B) showing results after subtraction of constants as indicated.

underscore the importance of high q data correction, and represent a rational method for assessing a minimum data range for shape reconstruction.

In order to test how constant subtraction affected shape determination, data from the above experiments were used as input for the programs GNOM and DAMMIF. DAMMIF runs produced a molecular mass determination from the data that was correlated to the constant subtracted similar to independent analysis [201]. Constant subtraction of 0.000-0.003 produced underestimates of molecular mass, while constants above 0.004 produced overestimates (Figure III.8A). Shape envelopes generated from this data also produced noticeable features correlating to constant subtraction (Figure III.8B). With too little of a constant subtracted, envelopes had large globular structures with irregular nodes protruding out. This is presumably an attempt to fit the data using an artificially small molecular mass. Corrections between 0.003 and 0.005 resulted in shape envelopes that matched the crystal structure well. Gross oversubtraction produced an enlarged shape envelope with diminished features. These constants and shape characteristics were consistent with results from another *ab initio* modeling program GASBOR, which automatically chooses a constant to subtract. GASBOR modeling with automatic subtraction of constants in the range 0.0055-0.0065 resulted in an overestimation of the size and washed out features. Upon deletion of the incorporated solvent layer before averaging, however, the expected envelope was obtained. Together, these observations help define the relationship between accurate solvent subtraction and shape envelope determination, and offer a qualitative and quantitative measure constant

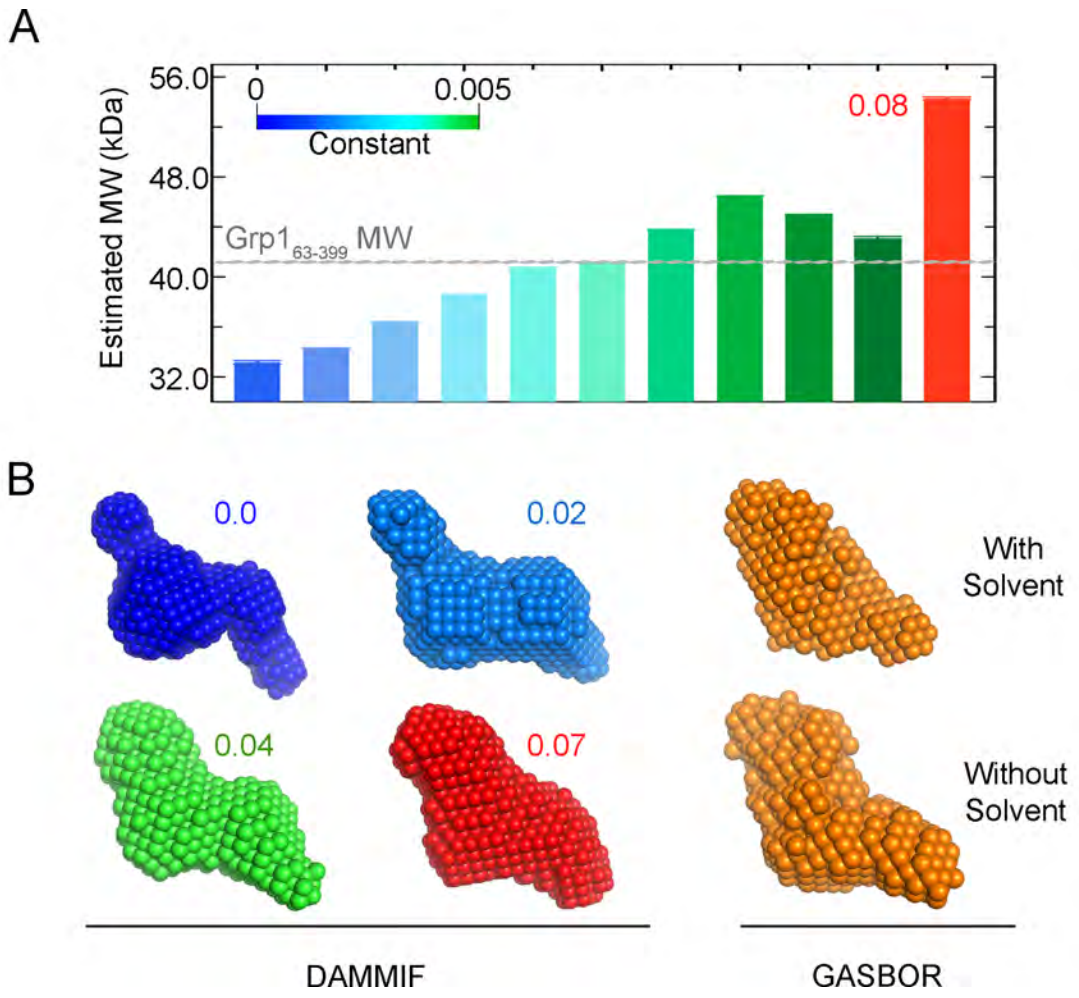

Figure III.8: Effects of Constant Subtraction on *ab initio* Shape Envelopes. A) Estimated molecular masses calculated from DAMMIF after constant subtraction as indicated. Corrected scattering intensities were normalized to 1.0 before analysis. Errors on select bars represent standard deviation from an arbitrary selection of 10 DAMMIF reconstructions. B, left) DAMMIF dummy atom envelopes after constant subtraction and envelope calculation from 100 DAMMIF cycles as described in Materials and Methods. B, right) Average GASBOR dummy atom envelopes from 10 cycles after averaging with or without added solvent molecules.

background subtractions of SAXS data in the high q region in the absence of WAXS data.

Discussion

SEC-SAXS offers the unique advantage of theoretically bypassing the most problematic aspects of data quality, producing buffer matched, aggregation-free, homogeneous samples at relatively low concentrations (1-5 mg/ml or less). Because of the increased sample quality and large amount of data, a detailed optimization of data processing and analysis methods could be performed. We have developed new processing and analysis techniques and tools, some of which may be widely applicable beyond SEC-SAXS. The Python-extensible software package DELA can be used to graphically and intuitively scout useable protein and solvent regions for standard buffer subtraction, and SVD analysis can inform on sample homogeneity over regions of interest in the SAXS chromatogram. Further, test cases for proteins with known structures suggest that linear combination of SVD components can be used to reconstruct protein scattering curves of sufficient quality for comparison with theoretical scattering as well as *ab initio* and rigid body modeling, provided the Guinier region is well determined and the range of data analyzed is homogenous with respect to protein composition.

While SEC-SAXS greatly oversamples the data, it presents challenges regarding buffer subtraction in that pure buffer regions are removed from those of the sample and may indeed contain different compositions before and after the sample depending on the

experiment. In these cases, oversampling can be exploited using SVD-LC, which allows indirect buffer subtraction from peak sample data alone. Although having slightly lower signal than when including buffer samples, SVD-LC analysis from a whole chromatography peak produced reasonable quality data, in part due to noise reduction inherent in SVD. This analysis will be very important with low-abundant samples, as it allows confident correction at low signals.

One potential strength of SEC-SAXS over other techniques that has not been fully explored is the ability to analyze multiple species in a single experiment. Many applications of SEC-SAXS are focused on obtaining structural information on hard to characterize samples due to poor stability or aggregation, which can be effectively mitigated by gel filtration. But even with well-behaving samples, there is a distinct advantage of SEC-SAXS as a method for easily and quantitatively studying homomeric or heteromeric complexes. Indeed, in the $Grp1_{63-399}$ 24mL data set, SEC-SAXS was able to differentiate between dimeric and monomeric forms of the protein in solution. Upon further analysis, it was reasoned that the dimer is a concentration-dependent species produced as an artifact of high concentrations (Figure III.6). It is not entirely clear why the I307E mutant shows increased dimer formation, but analysis of the crystallographic dimer led to the hypothesis that the acidic residue could increase local contacts between the PH domain and the Sec7 domain (Figure III.9).

When optimizing direct buffer subtraction, it was additionally puzzling that many of the same effects seen in traditional SAXS experiments such as ambiguous buffer subtraction results and interparticle interference were still seen in Guinier plots of SEC-

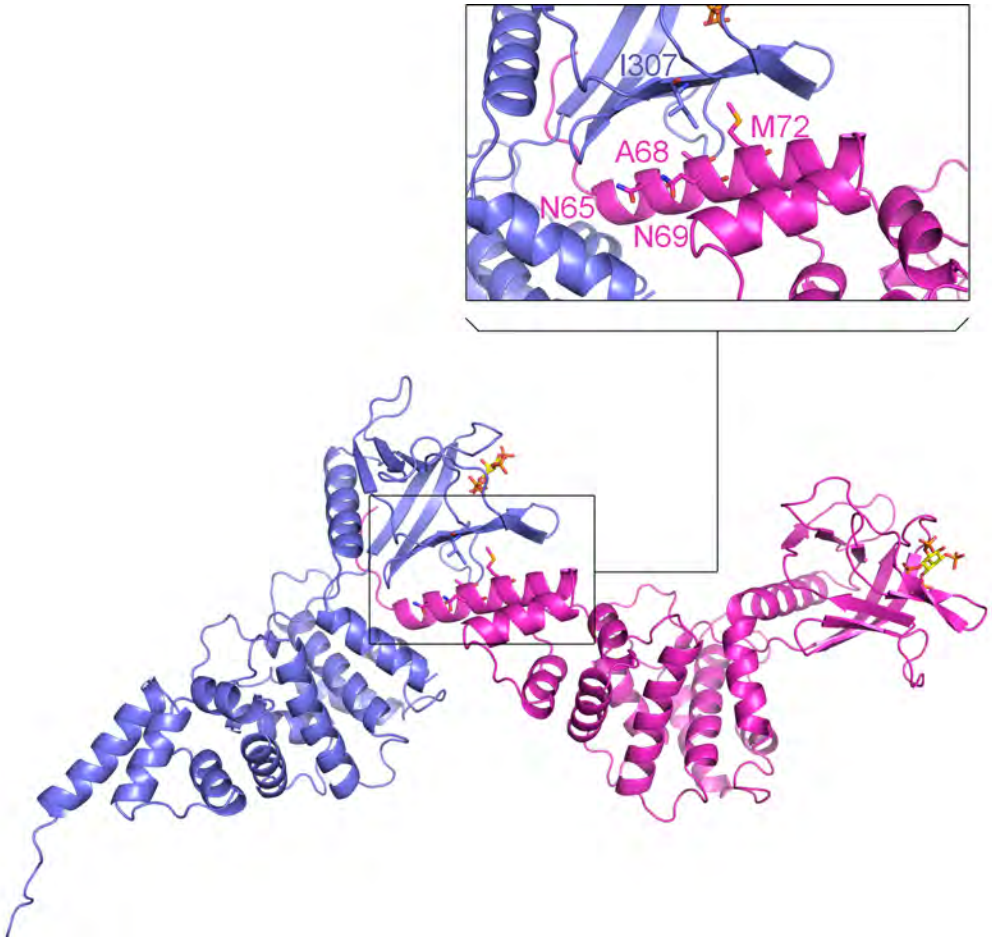

Figure III.9. The $Grp1_{63-399}$ Crystallographic Dimer. Dimer after redefinition of the asymmetric unit in the $Grp1_{63-399}$ crystal structure (PDB ID 2R09). Ile 307 and surrounding residues are shown in sticks.

SAXS data despite the high quality of the data and low concentration of samples unlikely to result in interparticle effects. These effects were also seen by overcorrection for buffer components in SVD-LC reconstruction (Figure III.5B), consistent with the hypothesis that these effects were manifested in excluded solvent volume at high concentrations that could be corrected in these data [201]. A direct buffer subtraction optimization was developed that explores a constant which, when applied to the protein scattering curve before buffer subtraction, produces the best R^2 value in a Guinier fit, similar to the optimization of linear combination after SVD. This accounts for excluded protein volume [226]. Thus buffer subtraction can also be efficiently optimized using a linear Guinier region when sample concentrations are accurately measured.

The results of this analysis represent new developments in the area of SEC-SAXS analysis that take advantage of the unique characteristics of these data, but that are applicable to all types of SAXS data to increase the breadth of macromolecular samples that are viable for this technique. We envision an application for these analysis techniques in complex experiments in which an accurate measure of the buffer sample is difficult or the components are at a low population and require merging of many low intensity scattering profiles.

Materials and Methods:

Cloning and Protein Purification

DNA constructs were amplified by polymerase chain reaction (PCR) with Vent polymerase (NEB), digested with BamH1 and SalI or XhoI, and ligated into a modified

pET15b vector. Mutations were generated using whole plasmid PCR with QuikSolution (Stratagene) and DpnI (NEB) digestion. After transformation of BL21(DE3) cells (Novagen) with plasmids, cells were grown in 2×YT with 100 mg/L ampicillin to an OD_{600} of 0.2 and expression induced with 50 μM IPTG at 18°C for 16 hrs. After harvest, cells were resuspended in buffer (50mM Tris pH 8.5 or 7.5, 150mM NaCl, 2mM $MgCl_2$, 0.05% 2-mercaptoethanol) and frozen at -80°C. Cells were thawed and incubated with 0.1 mM PMSF, 0.2 mg/ml lysozyme, and 0.01 mg/ml protease free DNAse I (Worthington). After sonication, lysates were supplemented with 0.5% Triton X-100 and centrifuged at 30,000 × g for 1 hr. Proteins were purified on batch Ni-NTA beads followed by ion exchange on HiTrap Q or SP and gel filtration on Superdex-75 or 200 (GE Healthcare).

Gel Filtration Analysis of $Grp1_{63-399}$ Dimers

Gel filtration was performed on a Superdex 200 10/300 (24 mL) column in a buffer consisting of 20 mM Tris pH 8.0, 150 mM NaCl, 2 mM $MgCl_2$, and 0.1% 2-mercaptoethanol. Samples were incubated at room temperature for 5 hours and 200 μL samples were loaded and eluted at 4° C. Standards were analyzed by gel filtration using the high molecular weight calibration kit (GE Healthcare).

SEC-SAXS

SEC-SAXS was performed at Bio-CAT, Sector 18-ID at the Advanced Photon Source (APS) in Argonne National Labs, Argonne, IL. 100 – 500 μl (10-20 mg/ml) or

100 µl (10 mg/ml) samples were loaded, respectively, onto 24 or 3 ml Superdex-200 columns (GE Healthcare) equilibrated with buffer (20 mM Tris pH 8.0, 150 mM NaCl, 2 mM $MgCl_2$, 0.1% 2-Mercaptoethanol (β-Me), and 1µM Inositol(1,3,4,5)tetrakisphosphate (IP_4)). $Grp1_{63-399}$ and Arf6 samples were incubated with 1.2 molar excess of IP_4 for 1-5 hrs before experimentation. Columns were connected in line with the sample capillary and data collected on a Mar CCD detector at a distance of 255.3 cm with 1 s exposures taken at 5 s intervals during elution. Data was monitored by UV chromatogram and showed consistent and clear separation of monomeric protein on both columns. Samples were calculated to be between 1-3 mg/ml at the peak of the protein elution profile.

Scattering Curve Calculations

Theoretical scattering curves from relevant crystal structures Grp1 (PBD 2R09) and Arf6 (PBD 2JFX) were generated using the program CRYSOL [213] with buffer subtraction.

Guinier Analysis

The Guinier plots are shown as log I(q) vs. q^2 plots and fits are based on the following approximation [227], [219] :

$$I(q) = I(0)e^{[-(q^2R_G^2)/3]}$$

Where I(0) is scattering intensity at q = 0, and R_g is the radius of gyration, which gives a measure of the overall size and shape of the protein. R_g is visualized as the slope of the linear region on a Guinier plot in a given q range. The Guinier region to fit was set such

that the $q*R_g$ was ≤ 1.3. Larger sizes, either through increases in molecular weight or elongated molecules, have steeper Guinier regions and larger R_g values.

P(r) Distribution Calculation

Pairwise Distribution Functions P(r) [228, 229] were calculated by indirect Fourier transform methods. . Distributions from the software package PRIMUS [219] were generally similar to Maximum Entropy Method (MEM) distributions calculated with DELA in test cases and were used as input for dummy atom modeling. PRIMUS integrates standalone programs such as AUTORG, AUTOGNOM, and GNOM to automatically calculate R_g based on the R^2 of a Guinier plot fit which AUTOGNOM uses to calculate a maximum distance (D_{max}) applied to the distribution [230]. An additional smoothing function and a zero distance constraint at the chosen D_{max} are also applied.

Pilot experiments to test the accuracy of P(r) distributions were performed by comparing experimental and theoretical data from the $Grp1_{63-399}$ construct, whose crystal structure contains a structured N-term His_6 tag identical to the experimental construct. AUTOGNOM results produced informative distributions maintaining peaks and shoulders, and experimental data as well as theoretical data had near identical P(r) distributions. However, when the data were not forced to zero at the chosen D_{max}, extra frequencies of distances remained, resulting in a tail extending out to ~100 Å. This tail was absent from theoretical data calculated without the tag, suggesting that at this size range, small features are represented in the data and the distance cutoff from AUTOGNOM is inadequate to represent them. We therefore chose an analysis method

that systematically increased the D_{max} until the distribution naturally came down to zero as has been documented for other systems [231].

Construction of Shape Envelopes Using SAXS

For DAMMIF modeling, P(r) distributions were used as input for 100 cycles of DAMMIF. Structures were grouped in sets of ten and average envelopes were generated using DAMSEL and DAMSUP, DAMAVER, and DAMFILT [213]. The ten filtered sets were again averaged and filtered. Within initial sets of ten, Normal Spatial Discrepancy (NSD) values ranged between 0.5 and 0.6 ± 0.01-0.07 and for averaged sets of ten NSD was 0.452 ± 0.005.

For GASBOR modeling, 10 cycles of GASBOR were performed with 344 dummy residues corresponding to residues 63-399 of Grp1 with an N-terminal His_6 tag, and 4-residues to approximate scattering from bound IP_4. For comparison with DAMMIF models, water molecules in the resulting structures were either maintained or removed before averaging with the DAMAVER suite as described above. Superposition of shape envelopes with representative crystal structures or similar envelopes was performed using SUPCOMB [232].

CHAPTER IV: LARGE SCALE REORGANIZATION AND DYNAMICS IN RELIEF OF CYTOHESIN AUTOINHIBITION

Abstract

In Cytohesin Guanine Nucleotide Exchange Factors (GEFs), autoinhibition of the catalytic Sec7 domain can be relieved through binding of Arf6-GTP to an allosteric site that includes the PH domain as well as proximal autoinhibitory elements consisting of the Sec7-PH domain linker and C-terminal helix. The previously determined crystal structure of Arf6-GTP in complex with a C-terminal Cytohesin-3 (Grp1) fragment lacking the Sec7 domain suggests that relief of autoinhibition is driven by competitive interactions involving large-scale conformational rearrangements that reposition the autoinhibitory elements into grooves at the Arf6-GTP/PH domain interface, thereby promoting accessibility of substrate to the catalytic site.

Although a Sec7-PH domain linker contributes to the autoinhibited and active conformations, it is unclear whether it adopts a stable or flexible structure upon Arf6 binding. In this chapter, small angle X-ray scattering (SAXS) is used to analyze the domain organization and flexibility of autoinhibited and active Cytohesins in solution. Autoinhibited Grp1 produced scattering data and *ab initio* shapes similar to the theoretical scattering for the crystal structure, and truncation of the autoinhibitory C-terminal helix produced subtle changes in the overall architecture. Active Grp1-Arf6 tandem constructs showed small bulk shape changes compared to autoinhibited forms. However, the shape envelopes were incompatible with the arrangement of domains in the autoinhibited Grp1 structure. Comparison of the active tandem data with theoretical models through Monte Carlo and rigid body modeling showed a requirement for alternate

conformations or flexibility in the Sec7-PH domain linker irrespective of the mobility of the Sec7 domain. These findings suggest movement of the Sec7 domain upon Arf6 binding into a potentially membrane proximal location primed for Arf-GDP binding.

Introduction

Autoregulation of Guanine Nucleotide Exchange Factors (GEFs) is a common theme in GTPase biology, and can be coordinated with membrane association of both GEF and substrate molecules for tight spatiotemporal control of cell signals [26, 87, 133, 138, 141, 142, 145]. GEFs are typically recruited to membranes through protein-lipid and protein-protein interactions [10, 78, 107, 152, 154], while GTPases often associate with membranes through insertion of lipid-modified elements. Catalytic output may be regulated by post-translational modification or binding events at or with membranes, in some cases leading to integration of membrane targeting and exchange activity [233, 234].

Arf GTPases are essential for cellular trafficking, migration, division, insulin sensitivity, and cancer [39, 235]. In the GDP bound state, the molecule is transiently associated with membranes through a myristoylated (myr) N-terminal amphipathic helix and is incompetent to bind classical effectors. Concerted association of the helix with membranes along with nucleotide exchange by Sec7 domain GEFs causes structural changes that allow stable membrane association of the N-terminus and effector binding to epitopes involving the canonical switch regions [59-61, 236, 237].

Cytohesin GEFs perform GDP to GTP exchange on Arf GTPases and include Cytohesin (Cyth1), ARNO (Cyth2), and Grp1 (Cyth3). They contain an N-terminal heptad repeat (HR) region, a catalytic Sec7 domain, and a Pleckstrin Homology (PH) domain. The Sec7 domain is physically autoinhibited by a Sec7-PH linker and a C-terminal helix including a polybasic region (pbr) that competes for Arf-GDP binding to

the Sec7 domain. Autoinhibition is relieved through phosphorylation of the C-terminal helix by protein kinase C (PKC) or binding of both phosphatidyl inositol (PtdIns) (3,4,5) tris-phosphate (PIP_3) to the PH domain as well as Arf6-GTP to the PH domain and flanking autoinhibitory elements [31, 133, 150]. Arf6-GTP binds the PH domain including the linker and C-terminal helix, sequestering the autoinhibitory elements into nascent grooves formed at the Arf6-GTP/PH domain interface and creating a large allosteric shift compared to the autoinhibited configuration. Arf6-GTP binding is primed (i.e. the affinity is enhanced by roughly an order of magnitude) by binding of the PH domain to PIP_3 or its head group inositol(1,3,4,5)tetrakisphosphate (IP_4). Relief of autoinhibition is also aided by membrane association of substrate and activating GTPases, further integrating the autoregulatory mechanism with membrane recruitment. [31, 151, 234]

Despite detailed insights into relief of autoinhibition by Arf6-GTP from structural and biochemical analysis, there is no structural information about the Sec7 domain in the Arf6-stimulated state, as the crystalized complex only includes the allosteric site [234]. The extent of mobility in the Sec7 domain upon binding to Arf6 is unknown. It is further unclear whether the conformation of the linker in the structure of the Arf6-GTP complex with the allosteric site is representative of the complex with the full length protein in solution, or an infrequently populated location stabilized by existing crystal contacts between molecules.

To understand the potentially complex Sec7 rearrangements upon Arf6 binding, Size Exclusion Chromatography in line with Small Angle X-ray Scattering (SEC-SAXS)

was performed with Cytohesin constructs in autoinhibited and active forms as well as Grp1-Arf6 fusion constructs. These showed a large scale rearrangement of Cytohesins upon activation compatible with relief of autoinhibition, but incompatible with structural observations of a fully ordered Sec7-PH domain linker. Monte Carlo ensemble modeling coupled with rigid body modeling indicate a requirement for greater flexibility in the N-terminal region of the linker than observed in the crystal structure of the Arf6-GTP complex. While this flexibility allowed multiple orientations of the Sec7 domain, the data require a location of the Sec7 domain greatly extended from the autoinhibited conformation. Together, these data suggest a mechanism for relief of autoinhibition in which Arf6 promotes a mobile, but more membrane proximal position of the Sec7 domain.

Results

SEC-SAXS of Autoinhibited and Active Grp1 Constructs

To investigate structural changes of Cytohesins upon relief of autoinhibition, SEC-SAXS combined with Singular Value Decomposition (SVD) and Linear Combination (LC) as described in Chapter III was performed on Grp1 constructs including the Sec7 domain, PH domain, and CtH but lacking the heptad repeat region. Proteins were concentrated and incubated at 10 mg/mL with excess of IP_4 and loaded onto a SEC column in line with SAXS in the presence of IP_4. Reduced data was normalized by the incident/transmitted beam intensities and the protein scattering reconstructed by linear combination of the two most significant components from SVD,

followed by standard Guinier and pairwise distribution function (PDDF, P(r)) analysis (Chapter III, and [218, 219]). P(r) distributions were calculated with GNOM and *ab initio* shape envelopes with DAMMIF and GASBOR [213, 217, 238]. Data were collected on known autoinhibited ($Grp1_{63-399}$) and fully active constructs ($Grp1_{63-390}$) [133].

As shown in Figure IV.1A for $Grp1_{63-399}$, total intensity chromatograms showed a predominately homogeneous protein peak that was reasonably well separated from minor contaminants. The reconstructed protein scattering for $Grp1_{63-399}$ was almost identical to the theoretical scattering calculated with CRYSOL (Figure IV.1B). Small deviations in the high q region likely reflect expected mobility of the histidine tag in solution compared with the ordered conformation trapped by intermolecular contacts in the crystal structure. The P(r) distribution for the $Grp1_{63-390}$ construct, with catalytic activity close to that of the isolated Sec7 domain [133], overlapped well with the autoinhibited data except for the appearance of a clear shoulder at about 50 Å (Figure IV.Ic). As expected, the $Grp1_{63-399}$ shape envelope closely resembled the crystal structure of this construct. Unexpectedly, the $Grp1_{63-390}$ shape envelope rather than showing limited or variable dummy atom density in modeling replicates indicative of greater flexibility, showed a consistent shape (NSD_{avg} = .601 ± 0.016) that was elongated compared with the envelope for $Grp1_{63-399}$ (Figure IV.1D). Thus, the scattering data in solution are consistent with the overall structure seen in the crystallographic studies of the autoinhibited state, and show an unambiguous shape envelope in the absence of the C-terminal helix, suggesting that

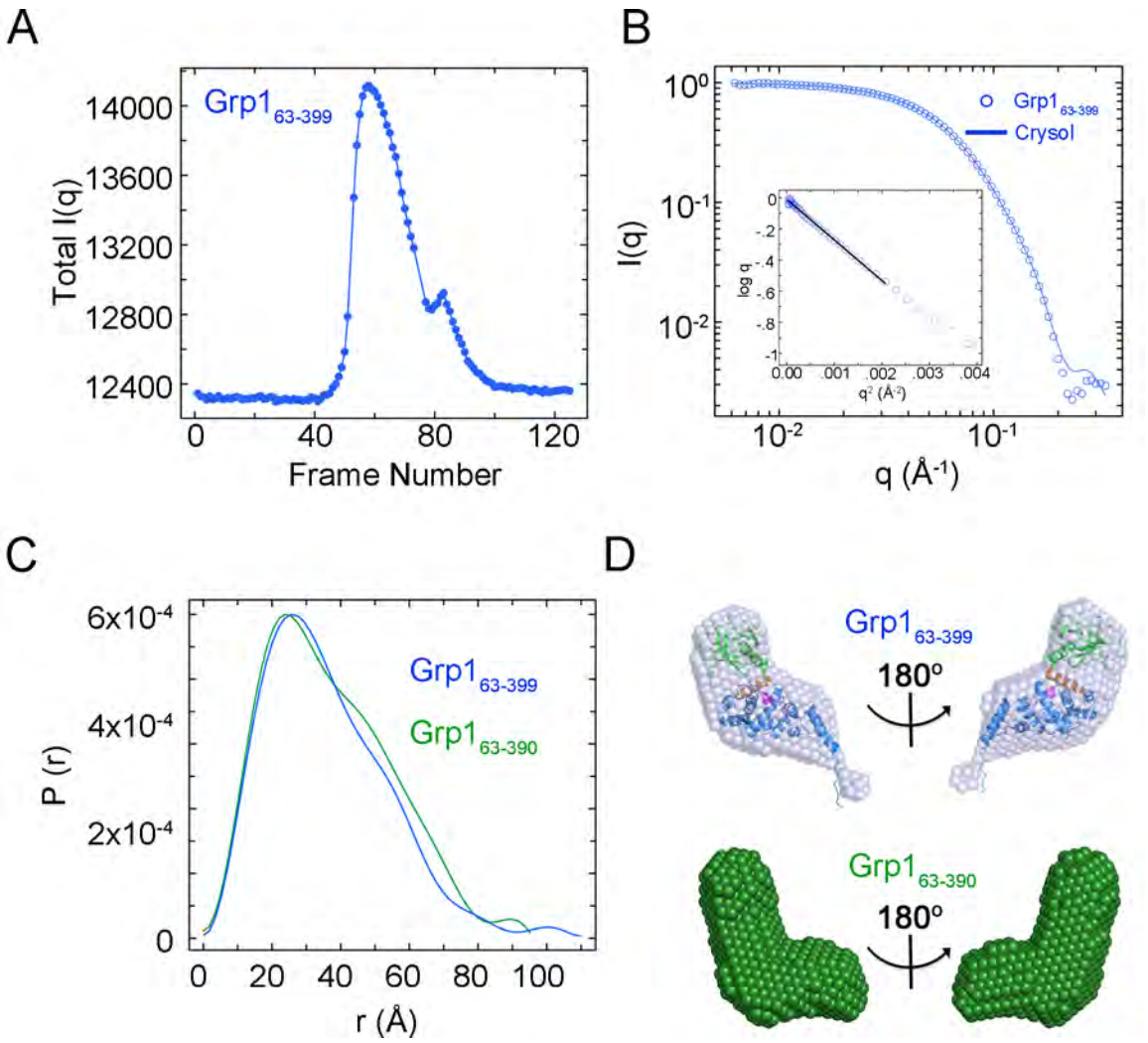

Figure IV.1: SEC-SAXS analysis of $Grp1_{63-399}$ and $Grp1_{63-390}$. A) Total Intensity Chromatogram of $Grp1_{63-399}$. B) Corrected scattering curves (circles) and theoretical scattering curves produced from CRYSOL (line). Inset is Guinier plot of the corrected data with fit. C) P(r) distributions of $Grp1_{63-399}$ and $Grp1_{63-390}$ constructs as indicated. D) *ab initio* shape envelops for $Grp1_{63-399}$ and $Grp1_{63-390}$ constructs.

the Sec7 domain may undergo a rearrangement to sample a less restrictive though perhaps still limited conformational space.

Relief of Autoinhibition in Grp1-Arf6 Tandems

Assessing structural changes in a Grp1-Arf6 complex in solution experiments is complicated by the high K_D (18 μM) [234] and simultaneous requirement for PIP_3 or IP_4 [31, 133, 151, 234]. These properties would lead to heterogeneous populations of bound and unbound states within the practical concentration range of SEC-SAXS, confounding analysis. To circumvent this problem, we designed a suite of Grp1-Arf6 tandem constructs for biochemical and SAXS analysis based on the crystal structure of the Grp1 PH domain with N-terminally truncated Arf6 [234]. Modeling in PyMOL and COOT indicated that a direct tandem construct could potentially allow relief of autoinhibition in cis due to the fortuitous 19 Å distance between the peripheral Grp1 C-terminus and the truncated Arf6 N-terminus in the crystal structure that could accommodate the Arf6 amphipathic helix without obvious steric conflicts. The presence or absence of a 6 Gly-Ser (6GS) linker between the Grp1 and Arf6 modules was also generated to provide a more flexible linkage and to assess uniqueness of overall kinetic and structural observations. These constructs were optionally supplemented with C-terminal histidine or SUMO tags to enhance the confidence of domain assignments, and generated with or without the Grp1 K340A mutation known to abolish Arf6-GTP dependent relief of autoinhibition and membrane recruitment [31, 217]. All constructs included an N-

terminal His_6 tag and, for Arf6, the Q67L mutation that strongly impairs the already low intrinsic rate of GTP hydrolysis (Figure IV.2A) [239].

To investigate autoinhibition in the context of tandem constructs, *in vitro* GEF assays for each construct were performed measuring mant-GDP exchange on N-terminally truncated Arf1-GDP (NΔ17Arf1) (Figure IV.2B). Assays using $Grp1_{63-399}$ in the presence of 80 μM N-terminally truncated Arf6 and IP_4 showed an 11-fold increase in catalytic efficiency (k_{cat}/K_M) over $Grp1_{63-399}$ alone. Assays with direct Grp1-Arf6 tandem constructs loaded with non-hydrolyzable GppNHp in the presence of IP_4 showed 4- to 5-fold increased activity, which was increased to 13- to 20-fold by addition of the 6 Gly-Ser (6GS) linker or C-terminal His_6 or SUMO tags. Removal of IP_4 from the system reduced stimulation in all cases, and the K340A substitution nearly abolished the effect.

Despite differences in stimulation among the suite of constructs, these assays indicate that in all cases the biochemical and structural requirements for relief of autoinhibition are met when Grp1 and Arf6 are expressed as a single polypeptide. Decreased kinetic activity in tandems lacking the 6GS linker may be a result of competition for substrate binding to the Sec7 domain by Arf6 binding or incomplete relief of autoinhibition. Regardless, results from the Grp1-6GS-Arf6 tandems show kinetic properties consistent with fully autoinhibited and active states.

Investigation of Grp1-Arf6 tandems by SEC-SAXS

To explore potential mechanisms for differences in tandem constructs, and assess the overall domain architecture of the Arf6 stimulated state of Grp1, SEC-SAXS

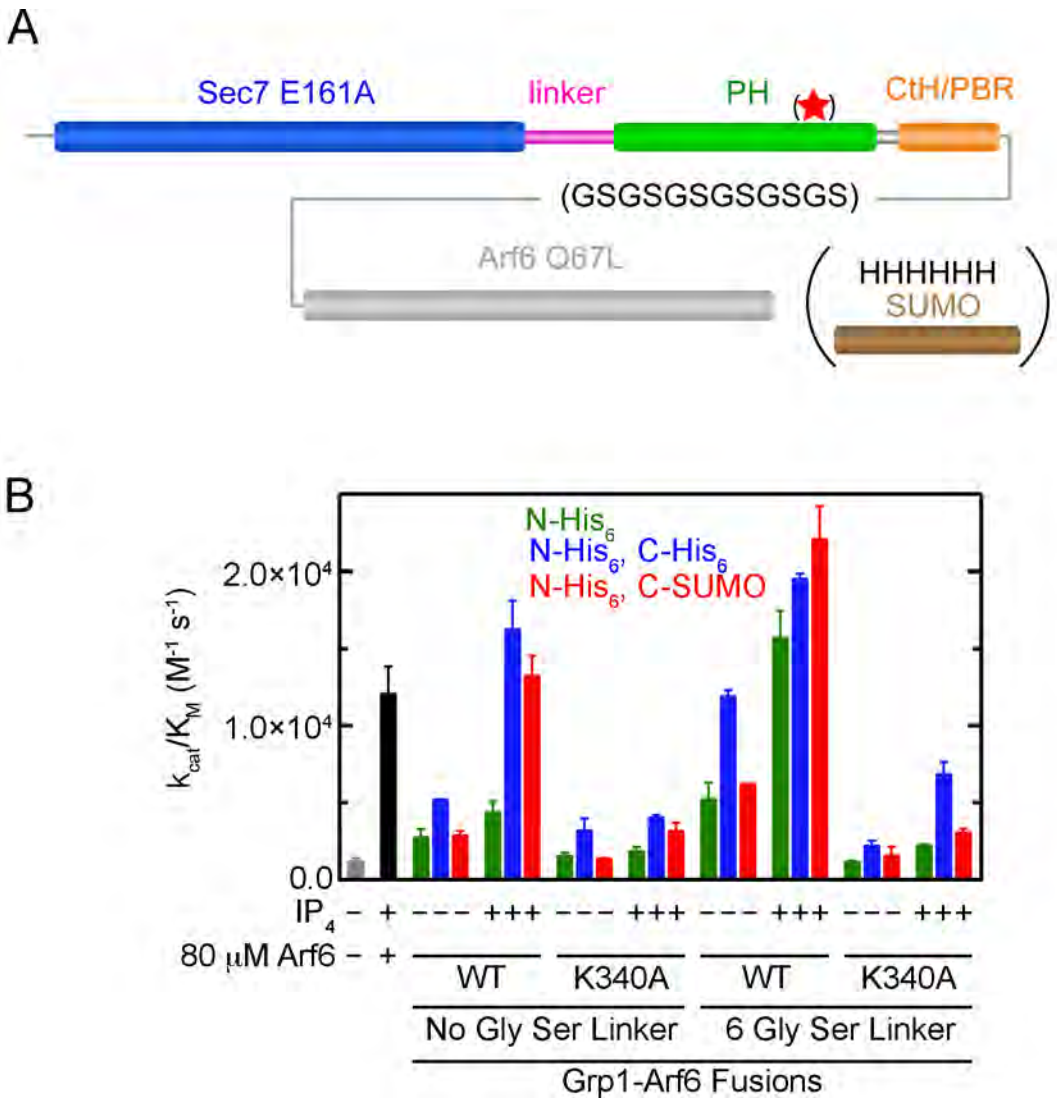

Figure IV.2: Domain Structure and Kinetic Analysis of Grp1-Arf6 Tandem Constructs. A) Cartoon of domain structure of constructs used for kinetic and SEC-SAXS analysis. Variable regions are in parentheses. Star represents the K340A mutation. B) Catalytic efficiency (k_{cat}/K_M) of Grp1-Arf6 tandem constructs for NΔ17Arf1-GDPmant exchange. Gray and black bars represent assays measuring kinetics with 125 nM Grp1 in the presence or absence of 80 μM NΔ13Arf6(Q67L) and 1 μM IP_4. All other assays represent assays using titrations of tandem constructs in the presence of absence of 10 μM IP_4.

experiments were performed in the presence of Ins(1,3,4,5)P_4 on Grp1-Arf6 tandems with or without the 6GS linker, K340A substitution, or C-terminal tags. These constructs also incorporate the E161A catalytically dead mutation in the Grp1 Sec7 domain to decrease intramolecular interaction of Arf6-GTP with the substrate site [89]. As shown in Figure IV.3 and Table IV.1, the Grp1-6GS-Arf6 tandem, which showed robust relief of autoinhibition in GEF assays produced high quality scattering data with an R_g of 32.5 Å and D_{max} of 136 Å. Similar high quality data was obtained for all other constructs.

The K340A variants had consistent increases in R_g and D_{max} values over wild type and clear shoulders in P(r) distributions (Figure IV.3B). These differences between WT and K340A constructs were small (< 2 Å difference in R_g) in comparison to the overall size of the construct but consistent between Grp1-Arf6 and/or Grp1-6GS-Arf6 constructs. C-terminally His_6 tagged constructs showed almost no changes in Rg values and P(r) distributions between K340A and WT tandems, likely due to increased overall flexibility. When comparing active constructs, addition of C-terminal His_6 or SUMO tags and/or deletion of the Gly Ser linker produced slight deviations in R_g and D_{max} values as well as broader P(r) distributions correlating with molecular mass differences between constructs (Figure IV.3C and D, and Table IV.1). Overall, however, these differences were small, and the overall P(r) distributions similar, consistent with a common overall domain organization and no major changes in shape or oligomeric state. Taken together, this data suggest slight changes in overall shape between autoinhibited and active constructs, and provide multiple data sets representing active structures for further analysis.

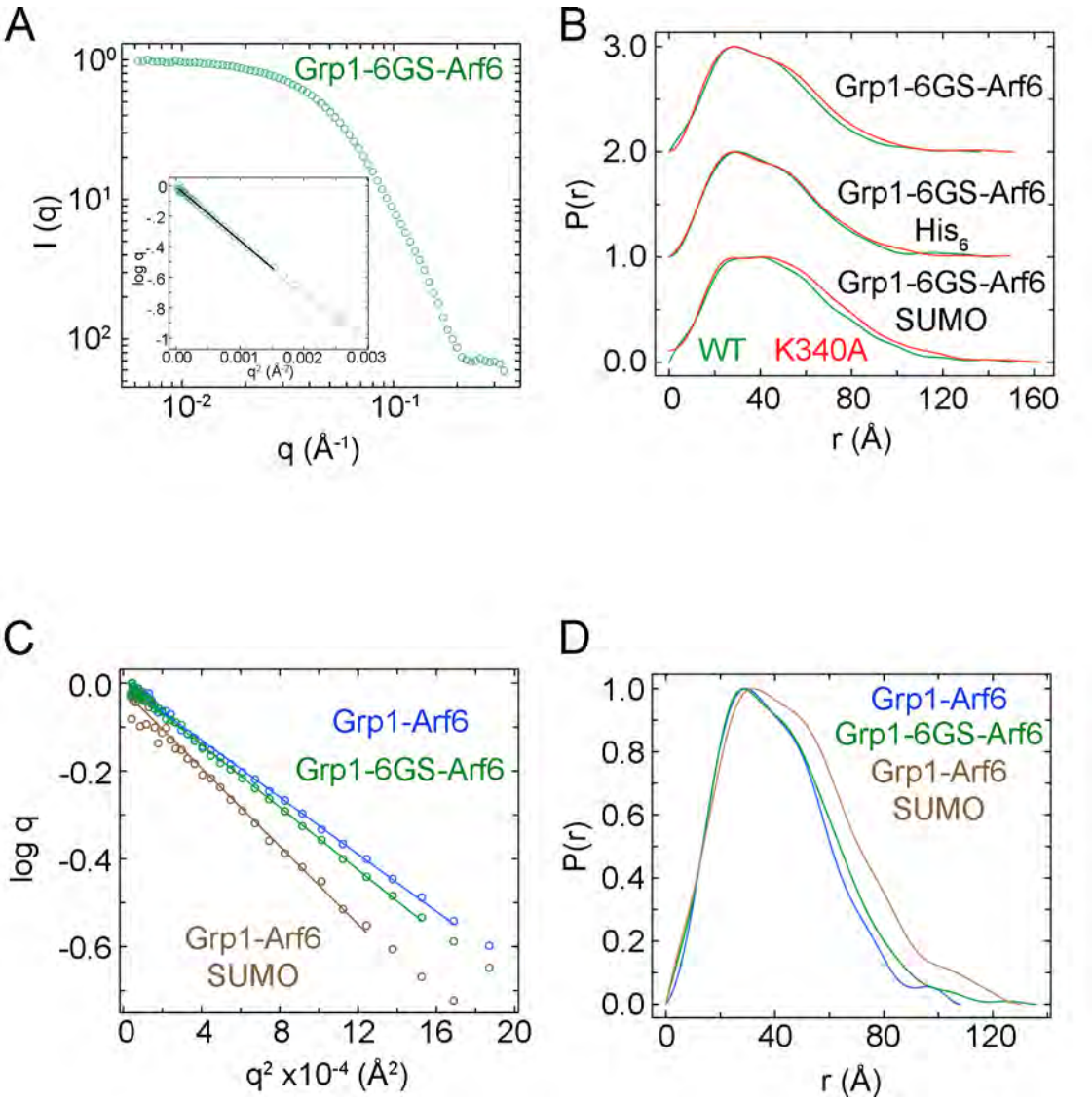

Figure IV.3: SEC-SAXS Analysis of Select Grp1-Arf6 Tandem Constructs. A) Corrected data and Guinier plot of the Grp1-6GS-Arf6 construct. B) P(r) distributions of wild type (WT) and mutant (K340A) Grp1-Arf6 tandems constructs C) Guinier plots of Grp1-Arf6 tandems including 6 Gly Ser and linkers and C-terminal SUMO tag for comparison. D) P(r) distributions of constructs in C.

	Rg		D_{max}
	Guinier	GNOM (s.d.)	
$Grp1_{63-399}$	27.7	28.3 (.103)	110
$Grp1_{63-390}$	27.6	28.8 (.175)	104
$Grp1_{63-390}$ (Dimer)	40.9	42.0 (.443)	140
His_6-Grp1-Arf6	30.9	31.6 (.082)	108
His_6-$Grp1_{K340A}$-Arf6	32.5	33.4 (.351)	134
His_6-Grp1-Arf6-His_6	32.1	32.4 (.112)	125
His_6-$Grp1_{K340A}$-Arf6-His_6	32.5	33.3 (.103)	132
His_6-Grp1-Arf6-SUMO	35.9	37.2 (.318)	130
His_6-$Grp1_{K340A}$-Arf6-SUMO	37.0	38.3 (.213)	165
His_6-Grp1-6GlySer-Arf6	32.5	33.4 (.162)	136
His_6-$Grp1_{K340A}$-6GlySer-Arf6	33.8	35.1 (.233)	151
His_6-Grp1-6GlySer-Arf6-His_6	33.5	35.0 (.183)	140
His_6-$Grp1_{K340A}$-6GlySer-Arf6-His_6	34.0	36.2 (.326)	182
His_6-Grp1-Arf6-6GlySer-SUMO	37.8	38.5 (.166)	151
His_6-$Grp1_{K340A}$-Arf6-6GlySer-SUMO	40.0	41.6 (.372)	187

Table IV.1. Radius of Gyration and D_{max} values from SEC-SAXS analysis.

Domain Architecture of Arf6-Stimulated Grp1 Tandems

In order to visualize the domain architecture for active Grp1-Arf6 tandems, dummy atom/residue modeling was used to create low resolution shape envelopes for the active Grp1-6GS-Arf6 constructs. Fusion constructs with or without C-terminal His_6 tags produced similar envelopes with a consistent size and shape (NSD between 0.453 and 0.535, Figure IV.4A and Table IV.1) in accordance with their similar P(r) distributions. A clear shape for the Grp1 PH domain-Arf6 complex and an additional extended lobe of dummy atom density was seen in all constructs, with only slight variability between constructs with or without C-terminal tags. The larger size and elongated shape of the Sec7 domain helped assign the domain into this portion of the shape envelope, although the orientation of the Sec7 domain could not be unambiguously deduced (Figure IV.4A). Determining the location of Arf6 in Grp1-Arf6 tandems was less straightforward than the Sec7 domain, owing to the similar sizes and shapes of the Grp1 PH domain and the Arf6 molecule at low resolution of the SAXS data. To resolve this, *ab initio* models for C-terminally SUMO tagged tandems were compared with constructs lacking the tag. The location of the extra dummy atom density for the SUMO domain was used to assign the PH and Arf6 portions of the envelope (Figure IV.4B).

The major domains without linking regions can be approximately positioned into the overall dummy atom envelope for further comparison (Figure IV.4C). Despite some ambiguity in the orientation of the domains and the low resolution of this shape information, these data suggest a stable average arrangement of the Grp1-Arf6 tandem architecture. Notably, the shape envelope from $Grp1_{63\text{-}390}$ construct showed clear overlap

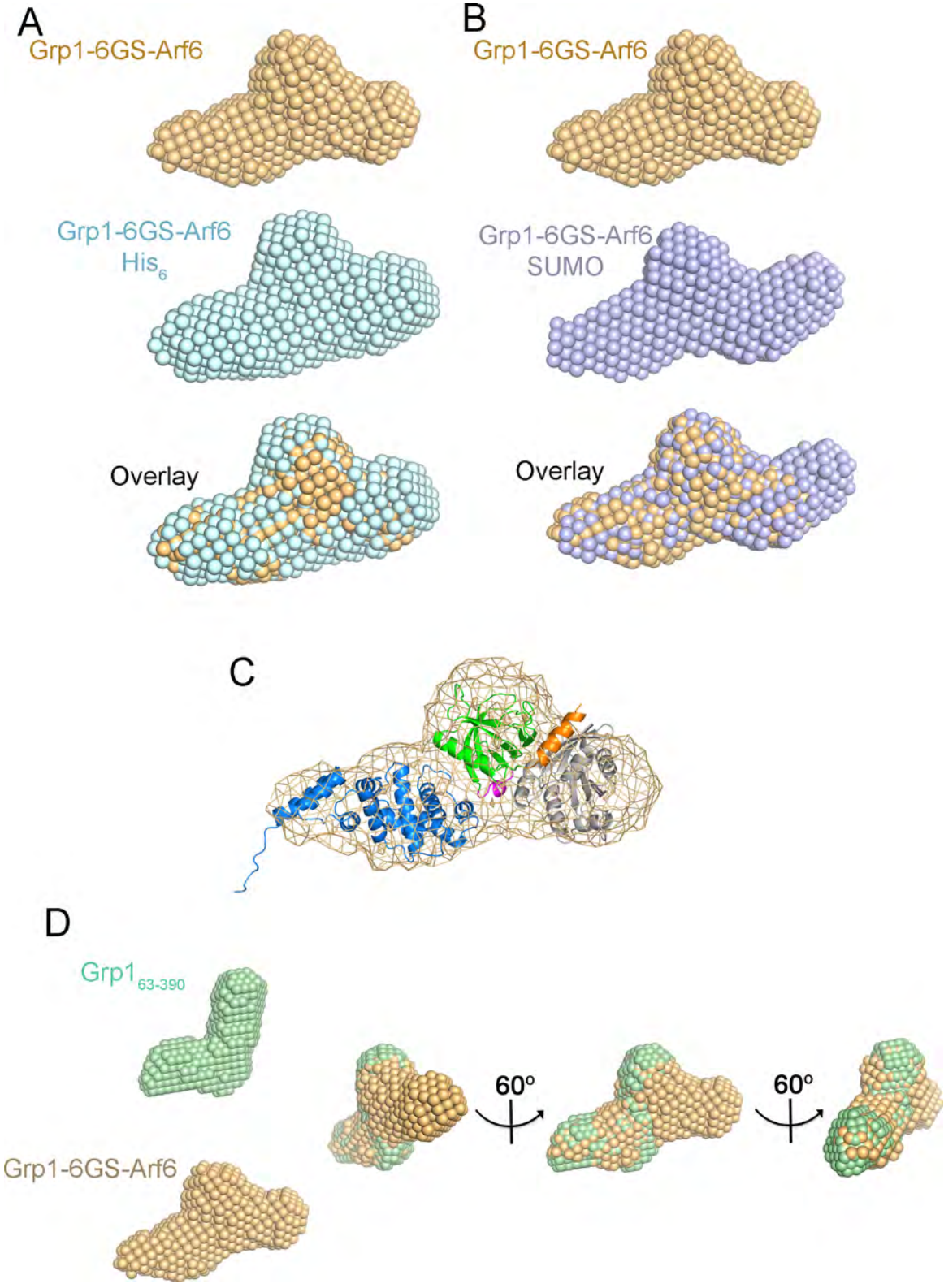
A
Grp1-6GS-Arf6
Grp1-6GS-Arf6
His6
Overlay
B
Grp1-6GS-Arf6
Grp1-6GS-Arf6
SUMO
Overlay
C
D
Grp1 63-390
Grp1-6GS-Arf6
60°
60°

Figure IV.4: Domain Assignment of Grp1-Arf6 Tandem Shape Envelopes. A and B) Filtered DAMMIF shape envelopes for Grp1-6GS-Arf6 (orange), Grp1-6GS-Arf6-His_6 (cyan), and Grp1-6GS-Arf6-SUMO (Blue). (A and B) Superpositions were performed with the program SUMCOMB. C) Manual domain assignment of Sec7 domain (from PDB 2R09) and PH-Arf6 complex (PDB 4KAX). D) Comparison of $Grp1_{63-390}$ and Grp1-6GS-Arf6 shapes showing overlap of the theoretical Sec7 and PH domains.

with the Sec7-PH domain portion of the Grp1-6GS-Arf6 envelope, consistent with similar active state organization between the Sec7 and PH domains. (Figure IV.4D).

Defining Structural Rearrangement in Grp1-Arf6 Tandem Constructs

The Grp1 Sec7-PH domain linker contributes to autoinhibition by interaction with the Arf1-GDP binding site [133]. Similarly, the Grp1-PH/Arf6 interface engages the linker in a nascent groove at the interface [234]. Superposition of the PH domain in the autoinhibited and Arf6-bound structures led to a hypothetical model for active Cytohesins whereby sequestration of the linker repositioned the Sec7 domain (Figure IV.5A, top). Alternatively, a hypothetical model assuming no rearrangement of the Sec7–PH linker upon Arf6 binding would result in a very different, but still compact architecture, as evidenced by the superposition of these models at the PH domain-Arf6 modules (Figure IV.5A, bottom).

To visually test these models with the *ab initio* shape envelopes from tandem constructs, the Grp1-6GS-Arf6 shape envelope was compared with these hypothetical structural models by docking the models into positions using the domain assignments in Figure IV.4. Not surprisingly, when the active Grp1-PH/Arf6 structure is docked into the active tandem shape envelope, the Sec7 domain density is oriented 90° in the opposite direction away from the hypothetical position (Figure IV.5B, top). Similarly, the Arf6 density is shifted approximately 90° from the location predicted when the autoinhibited Grp1 structure is anchored into the shape envelope at the Sec7 and PH domains (Figure IV.5B, bottom). As opposed to the compact structures theorized in from crystal

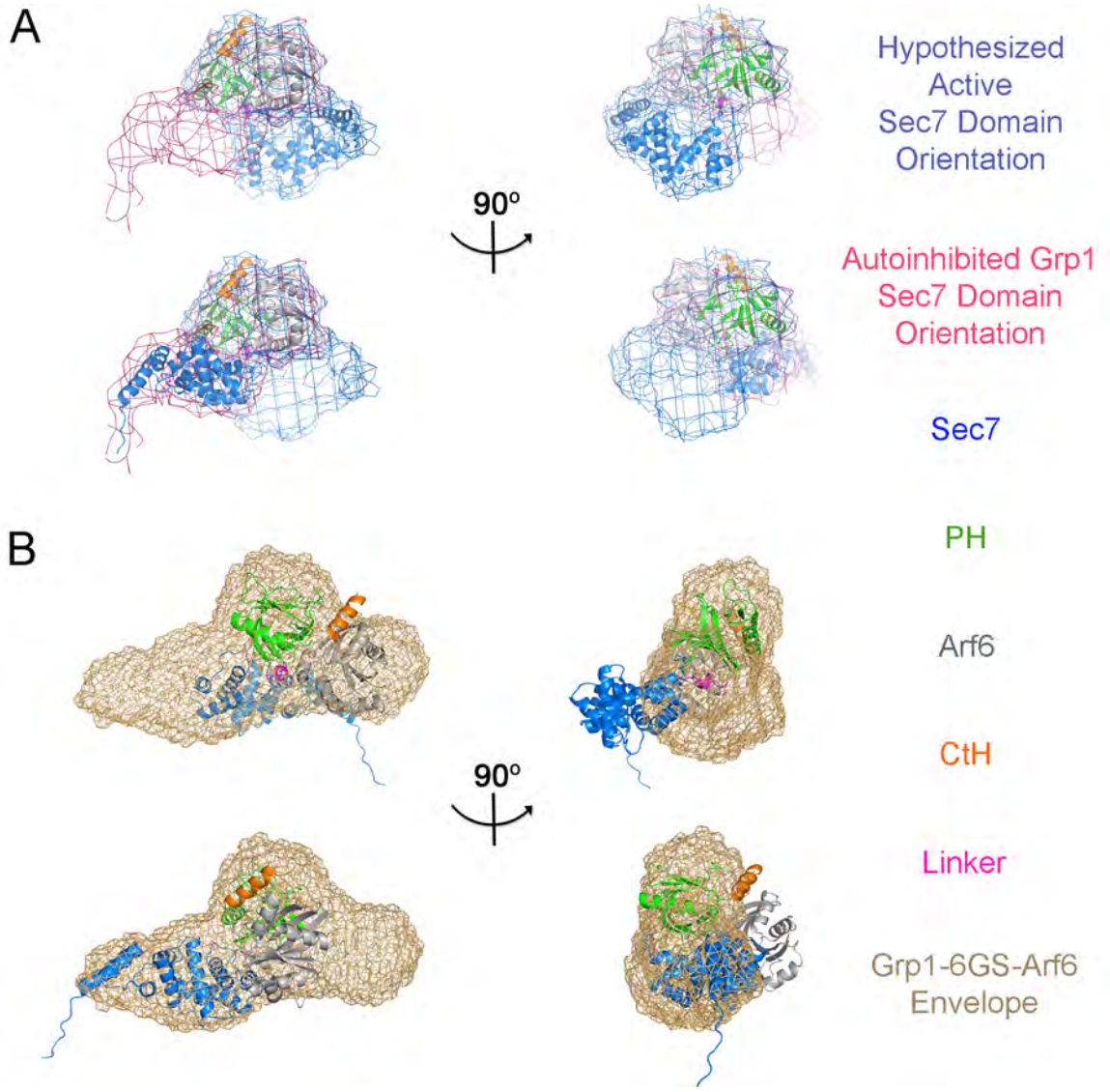

Figure IV.5: Structural Rearrangement in Active Grp1-Arf6 Tandem Constructs. A) Schematic representing theoretical autoinhibited (pink mesh) and active (blue mesh) conformations of the Grp1-Arf6 tandems. B) Comparison of Grp1-6GS-Arf6 shape envelope with conformations in A with (top) hypothetical active and (bottom) autoinhibited orientations of the Sec7-PH-Arf6 tandem.

structures, the observed architecture is elongated and planar. This comparison illustrates that at least part of the linker region most likely adopts an orientation different from that in the crystal structure of the Arf6-GTP complex. Because shape envelopes represent a low-resolution average of all conformations present in the sample, this could be explained by either flexibility in the linker region or large degrees of freedom in the Sec7 domain not represented in the final model. Manual modeling of the Sec7 domain and Grp1 PH-Arf6 structures into the envelope produced arbitrary or otherwise unsatisfactory results.

Flexibility of the Sec7-PH Domain Linker

The above observations led to a consideration of alternate conformations of the 15-residue Sec7-PH domain linker, which spans between two proline residues at position 251 and 265. In the autoinhibited state, residues 257-262 tightly bury into the Arf-GDP SWI binding surface [133]. Similarly, Leu 258 and Phe 262 bury into the interfacial groove within the Grp1 PH-Arf6 complex. However, this complex lacks clear density for residues 251 to 254 and, while the C-terminal helix contains mostly interactions with Arf6 in the Grp1-PH/Arf6 complex, residues 255 to 259 of the linker mediate interactions with the PH domain as well as Arf6 and contribute to crystal contacts. Consistent with these observations, truncations of linker residues up to residue 260 maintain full binding to Arf6-GTP [217].

In order to investigate the potential requirement for flexibility in the Sec7-PH domain linker, scattering data from Grp1-6GS-Arf6 constructs with single N-terminal

(Grp1-6GS-Arf6) or dual N/C-terminal His_6 tags (Grp1-6GS-Arf6-His_6) was analyzed directly through an approach using the Ensemble Optimization Method (EOM) [217, 240, 241]. In this process, 10,000 random structures were generated from a starting model with the Sec7 domain (residues 63-251) and the Grp1 PH–Arf6 complex treated as rigid bodies followed by a genetic algorithm to choose a best-fitting ensemble of structures. The beginning model for EOM was generated from arbitrary positions of the Sec7 domain and PH domain-Arf6 complex only slightly deviated from Figure IV.4C so as to maintain connectivity of the polypeptide chain and approximate the shape seen in *ab initio* envelopes. The C-terminal limit of the Sec7 domain was defined by a common structural feature in which a hydrophobic side chain caps the helical bundle at the C-terminus [26, 87, 89, 133], and the PH domain-Arf6 structure excluding the Sec7-PH domain linker was used as a model for these domains, using dummy residues to represent the N—terminal helix of Arf6 and the 6GS linker.

To assess flexibility requirements in the linker, the PH domain structure bound to Arf6 was used and the linker portion modeled as flexible dummy residues from residue 251 to residue 255 (the entire crystalized linker intact), 260 (retains minimal allosteric binding site for Arf6), or 265 (entire linker replaced with dummy residues). The 255 truncation mimics the compact active structure hypothesized in Figure IV.5A and the 265 truncation confers maximal flexibility. In all cases, dummy residues were used to represent the absent His_6 tags, the Sec7-PH domain linker as specified above, and the 6GS linker between Grp1 and Arf6. As the N-terminal portion of Arf6 is a predicted helix and the Sec7-PH domain linker is ordered in the Arf6-bound structure [234], this

model potentially overestimates the flexibility, but importantly allows for an unbiased sampling of flexibility requirements for the Sec7-PH domain linker.

EOM ensemble R_g distributions can be compared with the pool distribution to give a measure of experimental versus theoretical flexibility. For both Grp1-6GS-Arf6 and Grp1-6GS-Arf6-His_6, the ensemble produced from the 255 truncation model failed to fit the data (χ > 5.0 and 4.4, respectively), producing a distribution shifted far away from the pool that could not represent the experimental R_g values of 32.5 Å and 33.5 Å, respectively (Figure IV.6A). The ensemble distributions for the 260 truncation linker length fit the data (chi = 1.8 and 1.9) and were sharper and shifted to the right compared to the pool distribution, suggesting more elongated, overall structures than represented by random mobility of the domains (Figure IV.6B). Slightly better fits were obtained when the entire linker was mobile in the 265 truncation (chi = 1.35, and 1.43, respectively Figure IV.6C). Fits were generally better for the singly His_6 tagged construct in all cases, both in chi values and by visual analysis of the residuals (Figure IV.6D and E), but overall, the two constructs performed equally well in this analysis, offering an independent observation of structural relationships.

Refinement of Representative Structures Using Rigid Body Modeling

To further assess the conformational space potentially represented by the data, representative structures from EOM ensembles were used as starting models for rigid body modeling in CORAL [242], which uses simulated annealing to refine individual starting structures to the data. As shown in the example in Figure IV.7A, starting

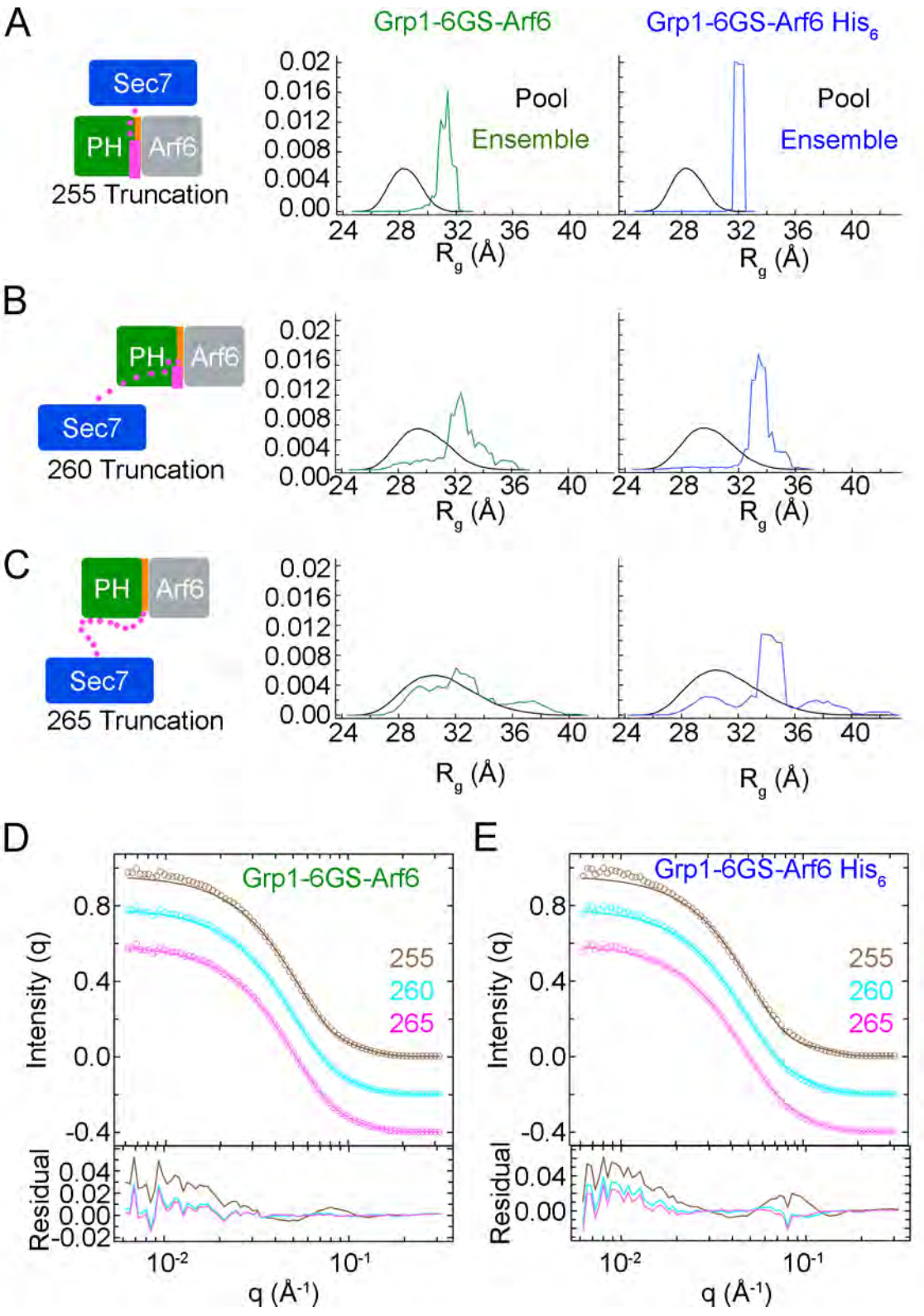
A
Grp1-6GS-Arf6
Grp1-6GS-Arf6 His6
Sec7
PH
Arf6
255 Truncation
Pool
Ensemble
Rg (Å)
B
260 Truncation
C
265 Truncation
D
E
Intensity (q)
Residual
q (Å-1)
255
260
265

Figure IV.6: Ensemble Optimization Method Analysis of Grp1-Arf6 Tandems with Flexible Linkers. Radius of gyration (R_g) distributions from pool (black) and best-fitting ensemble of structures from Ensemble Optimization Method (EOM) fit to the His_6-Grp1-6GS-Arf6 (green. left) and His_6-Grp1-6GS-His_6 (blue, right) data. Starting models in which Grp1 residues 252-254 (A), 252-260 (B), or 252:265 (C) were kept as flexible dummy residues. D) EOM ensemble scattering intensities (lines) fit to the data (circles) with residuals for the Grp1-6GS-Arf6 data set and models as indicated. E) As with D, but fitting to the Grp1-6GS-Arf6-His_6 data set.

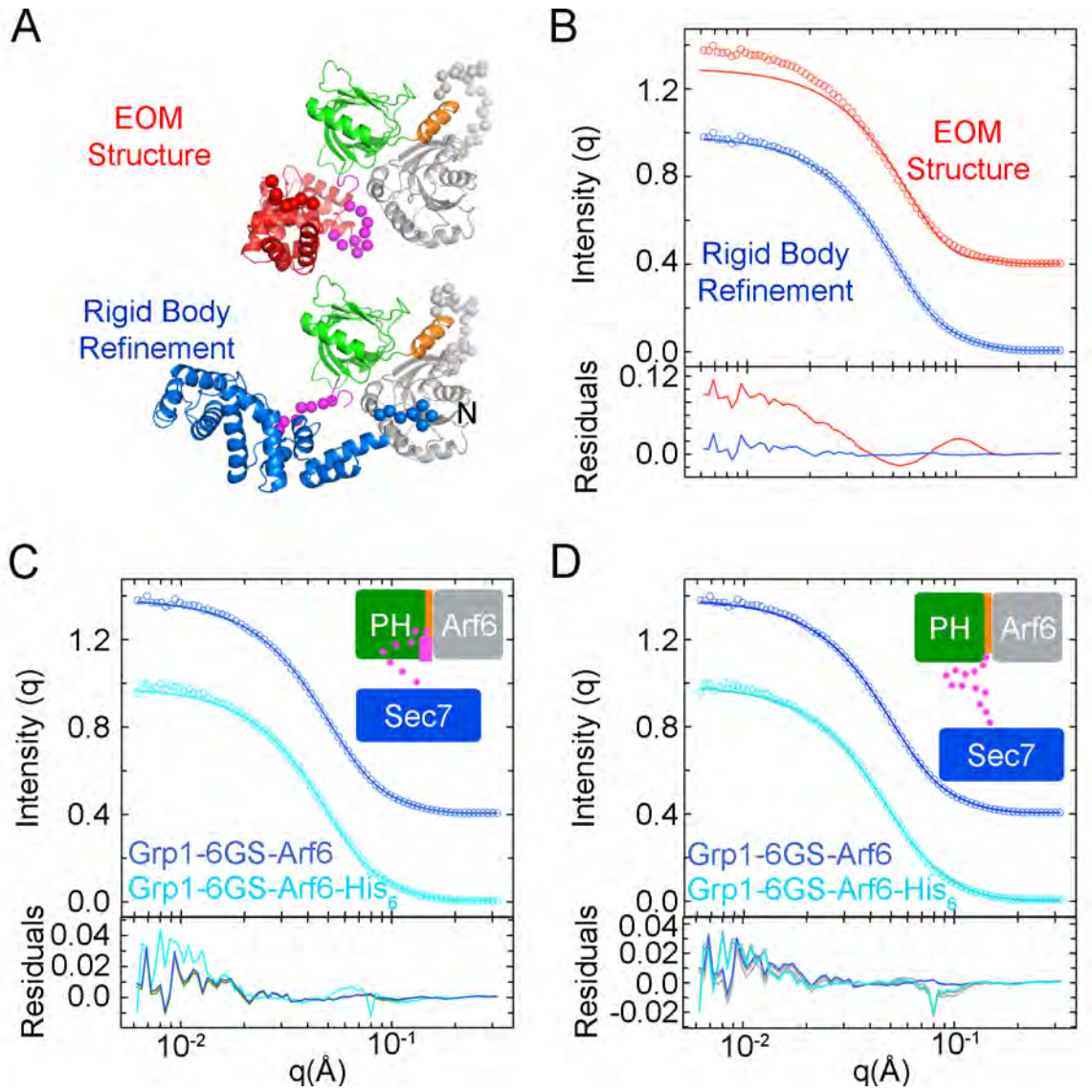

Figure IV.7: Rigid Body Modeling of Ensemble Optimization Method Structures for Grp1-Arf6 Tandems. A) Example refinement of representative Sec7 domain structure from Ensemble Optimization Method (red) using the Grp1-6GS-Arf6 tandem using Rigid Body Modeling in CORAL (blue). B) Fits of structures in A to the data. The fit for the EOM Structure was calculated in CRYSOL after constant subtraction. C) Fits from CORAL refinement of EOM structures with the 260 truncation of the Grp1 PH domain described in Figure IV.6. Representative models are for Grp1-6GS-Arf6 (blue) and Grp1-6GS-Arf6-His_6 (cyan), with others color in gray. D) Fits from CORAL refinement as in C, but using truncation of the Grp1 PH domain at residue 265.

structures that did not initially fit the data could be refined to a structure with deviations from the model equal to EOM results. Using such a wide variety of starting structures allows a relative measure of the uniqueness of any one result for a particular data set. Rigid Body Modeling was performed using the representative EOM structures from models truncating the PH domain at residue 260 or 265. Similar to the EOM modeling, the Grp1 PH domain-Arf6 complex was fixed during simulation. With the 260 truncation, refinement of EOM structures from both Grp1-6GS-Arf6 and Grp1-6GS-Arf6-His_6 data sets produced reasonable fits (chi = 1.8-2.0 from four structures, Figure 7b). Similar to relationships in EOM, starting models truncating the linker at residue 265 produced slightly better fits, with chi values of 1.8-1.9 and 1.4-1.6 for the Grp1-6GS-Arf6 and Grp1-6GS-Arf6-His_6 data, respectively (Figure IV.7C and D). These data confirm that some flexibility of the Sec7-PH domain linker is necessary to account for the Grp1-Arf6 tandem SAXS data, and provide reasonable starting models for further consideration.

To define possible models for the orientation of the Sec7 domain in these molecules, the refined structures from the EOM-CORAL analysis were compared. Representative examples of the structures are shown in Figure IV.8. When fitting the 260 truncation model to data from Grp1-6GS-Arf6 constructs with or without the His_6 tag, a variety of Sec7 positions were allowed that oriented the N-terminus toward the PH-domain/Arf6 complex (Figure IV.8A). However, the structures from refinement of the 265 truncation models produced structures with a much greater degree of freedom for the Sec7 domain resulting in variable N-terminal orientations (Figure IV.8B). As an

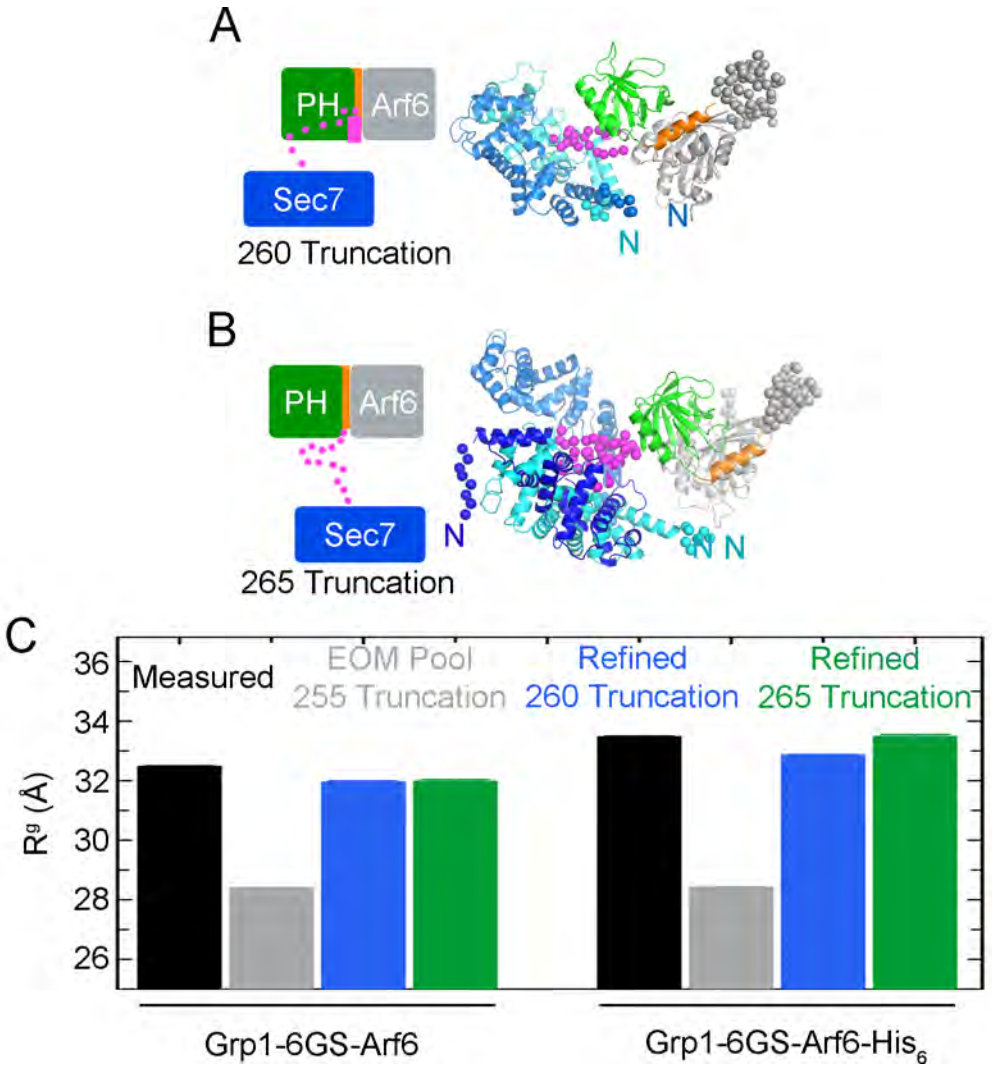

Figure IV.8: Orientations of the Grp1 Sec7 Domain after Rigid Body Refinement with Flexible Sec7-PH Domain Linker. A) Representative structures from EOM and CORAL refinement using either the Grp1-6GS-Arf6 or Grp1-6GS-Arf6-His_6 data with truncation of the Grp1 PH domain at residue 260. B) Representative structures as in A with truncation of the Grp1 PH domain at residue 265. C) Comparison of measured and calculated R_g values from EOM or EOM with CORAL refinement colored as indicated. For refined structures, R_g values were calculated using CRYSOL with constant subtraction. Error bars represent standard deviations of results in cases with more than one refined structure.

confirmation of this requirement, R_g values for refined structures in the 260 or 265 truncation models were in good agreement with the experimental values, but deviated from the average EOM pool R_g from the 255 truncation, which has a fixed linker position and randomized Sec7 position (Figure IV.8C.) Together, these observations underscore the necessity for flexibility and/or alternate conformations of the Sec7-PH domain linker from fully ordered conformation seen in the Arf6-bound crystal structure in addition to a likely dynamic Sec7 domain orientation to account for observed SAXS data.

Discussion

Autoregulation in GEFs presents a challenge for structural studies in that active forms often include flexible conformations that either preclude or are not adequately represented by crystallographic studies. For this reason, earlier studies of the active states of Cytohesins focused on biochemical determinants for relief of autoinhibition and structural studies in the absence of the catalytic domain. The work presented here represents the first structural description of active Cytohesins. SEC-SAXS analysis of Cytohesins stimulated through either truncation of the polybasic region or binding of Arf6-GTP showed subtle changes in overall structure compared to autoinhibited forms. *Ab initio* shape envelopes of these constructs showed an elongated architecture that appeared to be incompatible with a fully ordered Sec7-PH as observed in the crystal structure of the Arf6-GTP complex with a Grp1 construct lacking the Sec7 domain. EOM and rigid body analyses confirmed that dynamic flexibility, or at a minimum alternate conformations, for the N-terminal portion of the linker are necessary to account for the data. Although individual structures obtained from CORAL were able to fit the

data nearly as well as the optimized ensembles from EOM, the lack of interdomain contacts involving the Sec7 domain in the individual CORAL refined structures is consistent with a more dynamic view of the domain organization and, in particular the orientation of the Sec7 domain, such as that provided by the optimized ensembles.

Molecular determinants for relief of autoinhibition were maintained in Grp1-Arf6 tandems and overall *ab initio* shape envelopes were similar in SEC-SAXS experiments with or without the 6GS linker or C-terminal tags. Still, direct tandems showed relatively weak relief of autoinhibition that was increased by addition of the 6GS linker or C-terminal tags. This may be due to the necessary absence of the Grp1 E161A mutation in kinetic assays. When the catalytic Sec7 domain glutamate is present, the Arf6 in tandem may effectively compete for substrate binding. This would manifest more in constricted contexts, and less so when more flexibility is allowed within the Grp1-Arf6 linker or if a C-terminal tag hindered such an interaction. In SAXS experiments, the Grp1 E161A mutation blocks this interaction and favors Arf6 binding with the PH domain predicting a more consistent structure between the different active constructs.

Despite the resolution limits of SAXS, EOM and rigid body modeling helped to clarify a requirement for flexibility in the linker region of active Cytohesins. Given the well defined structures of the Sec7, PH, and Arf6 modules, flexibility would reasonably come from the autoinhibitory elements. In the Arf6-bound state, flexibility is unlikely in C-terminal portion of the Sec7-PH linker or the C-terminal helix due to extensive contacts of Phe 262 in the linker and multiple contacts involving the polybasic region in the Arf6 binding interface. Although N-terminal portions of the Sec7-PH linker are

located at the interface, the ordered conformation in the crystal structure appears to be stabilized via intermolecular contacts and there is evidently more flexibility in solution and/or in the context of constructs that contain the Sec7 domain. Consistent with these conclusions, defects in Arf6 binding were seen upon truncation of the linker at residue 263 but not 260 [234]. EOM served as a good test of these properties, since it uses a Monte Carlo approach and does not include assumptions about the position of the Sec7 domain. If the SAXS data only required mobility of the Sec7 domain, an active model simply using the known crystal structures would have fit the data. Instead, EOM and rigid body modeling results indicated that replacing the ordered Sec7 linker in the active model with flexible dummy residues up to residue 260 provides sufficient flexibility to fit the data, while the inclusion of the fully ordered linker disallowed elongated structures suggested by ab initio SAXS envelopes.

Because of resolution limits of SAXS and uncertainty about flexibility of the Sec7-PH domain linker, the exact orientation of the Sec7 domain is unclear and likely dynamic. As shown by EOM and rigid body modeling, models that include flexibility starting at residue 260 tend to be more elongated than the pool distribution, suggesting a more limited range of motion for the Sec7 domain. However, models with an entirely flexible Sec7-PH linker show a wide range of Sec7 domain orientations and positions. Despite this ambiguity, the Sec7 domain is shifted greatly from the autoinhibited state and in putative locations diverted from that hypothesized from a rigid linker region. This can be incorporated into a model for relief of autoinhibition integrating the current data, which can be constructed by simple superposition of well-fitting rigid body models with

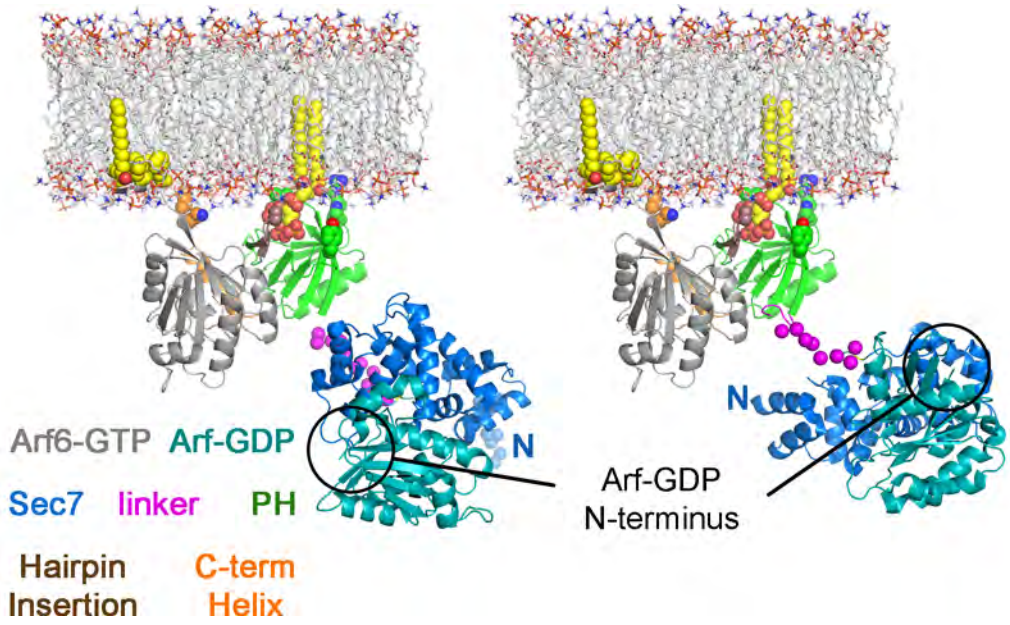

Figure IV.9: Model for Relief of Autoinhibition With Sec7 Domain Reorientation. Model was constructed using EOM-CORAL analysis of the Grp1-6GS-Arf6 data using 265 truncation (left) or 260 truncation (right) of the Grp1 PH domain. The model was superposed with the Grp1 PH-Arf6 crystal structure (PDB ID 4KAX). The Arf1-GDP/ARNO structure was superimposed at the Sec7 domain and Arf6-GDP aligned with Arf1-GDP for the location of the substrate Arf. The myristoyl and diacyl glycerol moieties on Arf6-GTP are modeled in configurations compatible with membrane insertion. The POPC bilayer membrane was derived from the coordinates of a molecular dynamics simulation [186].

Grp1 PH domain-Arf6 crystal structure and docking on a simulated membrane environment via the 1-phosphate of Grp1 similar to previous studies [234] (Figure IV.9). Strict orientation of the Sec7 domain N-terminus away from the Grp1 PH domain could orient the Arf-GDP binding site away from the membrane environment too far for membrane associated nucleotide exchange [60, 61, 77]. A more dynamic Sec7 domain and/or flexible linker may allow sampling of alternate Sec7 domain orientations suited for catalysis of membrane bound Arf-GDP. Finally, flexibility in the N-terminal portion of the linker in solution does not exclude the possibility that membrane association would further influence the orientation of the Sec7 domain.

It remains to be seen what role rearrangement of the Sec7 domain and or substrate binding plays in the overall mechanism integrating membrane targeting with autoregulation. Arf-GDP binding may help stabilize a membrane-proximal conformation, providing a concerted mechanism to create stable microenvironments of specialized GEF activity, as has been seen for BRAG2 [87]. In addition, Arf6-bound conformations of the Sec7 domain may itself contribute to membrane association through direct membrane insertion or dimerization of the heptad repeat regions. Further biochemical and structural studies with dimeric and substrate-bound Cytohesins could help to define the sequential or concerted mechanisms for relief of autoinhibition and membrane recruitment by GTP-bound Arfs and Arls.

Materials and Methods

Cloning and Protein Purification

DNA constructs were amplified using PCR with Vent polymerase (NEB), digested with BamH1, and SalI/XhoI and ligated into a modified pET15b vector. Mutants were generated using the whole plasmid PCR with QuikSolution (Stratagene) and DpnI (NEB) digestion. After transformation of BL21(DE3) cells (Novagen) with plasmids, cells were grown in 2×YT with 100 mg/L ampicillin to an OD_{600} of 1.0-1.2 and expression induced with 1mM IPTG at 37°C for 3 hrs. After harvest, cells were resuspended in buffer (50 mM Tris pH 8.5 or 8.8, 150 mM NaCl, 2 mM $MgCl_2$, 0.05% 2-mercaptoethanol) and frozen at -80°C. Cells were thawed and incubated with 0.1 mM PMSF, 0.2 mg/ml lysozyme, and 0.01 mg/ml protease free DNAse I (Worthington). After sonication, lysates were supplemented with 0.5% Triton X-100 and centrifuged at 30,000 × g for 1 hr. Proteins were purified on batch Ni-NTA beads followed by ion exchange on HiTrap Q or SP and gel filtration on Superdex-75 or 200 (GE Healthcare).

Nucleotide Loading

Tandem Grp1-Arf6 samples were incubated overnight at 4°C in 20 mM Tris, pH 8.0, 150 mM NaCl, 5 mM EDTA, and 1 mM DTT with a 10 fold molar excess of GppNHp. $MgCl_2$ was added to a final concentration of 10 mM and excess nucleotide removed by gel filtration on Superdex-75 (GE Healthcare) or using a D-Salt column (Pierce ThermoFisher).

GEF Assays

The fluorescence decrease upon mantGDP dissociation from Arf1NΔ17 was monitored in 20 mM Tris, pH 8.0, 150 mM NaCl, 2 mM $MgCl_2$ and 250 μM or 1 mM GppNHp. Grp1 or Grp1-Arf6 tandem constructs with or without active Arf GTPases and/or head groups were formatted into 96 well half area microplates (Corning) and incubated for 16-24 hrs at 25°C. Reactions were initiated by addition of 1 μM Arf1NΔ17-mantGDP and monitored using a Safire microplate spectrophotometer (Tecan) with excitation and emission wavelengths of 360 nm and 440 nm, respectively. Observed rates constants (k_{obs}) were obtained by fitting with $I_t = (I_0 - I_\infty) \exp(-k_{obs}\, t) + I_\infty$, where I_t, I_0, and I_∞ are the emission intensities at times t, t = 0, and as $t \rightarrow \infty$. Catalytic efficiencies (k_{cat}/K_M) were obtained from the slope of a linear least squares fit with $k_{obs} = (k_{cat}/K_M)[GEF] + k_{intr}$, where k_{intr} is the intrinsic dissociation rate constant.

SEC-SAXS

SEC-SAXS was performed at Bio-CAT, Sector 18-ID at Advanced Photon Source (APS) in Argonne National Labs, Argonne, IL. 100 – 500 μl (10-20 mg/ml) or 100μl (10 mg/ml) samples were loaded onto 24 or 3 ml Superdex-200 columns, respectively (GE Healthcare) equilibrated with buffer (20 mM Tris pH 8.0, 150 mM NaCl, 2 mM $MgCl_2$, 0.1% 2-Mercaptoethanol, and 1μM inositol(1,3,4,5)tetrakisphosphate (IP_4)). Columns were connected in line with the flow cell for SAXS data collection and 1 s exposures were taken at 5 s intervals during elution. All samples containing Grp1 either alone or in tandem were incubated with 1.2 molar

excess of IP_4 for 1-5 hours prior to injection. The UV absorbance at 280nm was monitored during chromatography and showed consistent and clear separation of monomeric protein on both columns.

Data Processing

A detailed description of data processing including method development is given in Chapter III. Briefly, data images were radially averaged over the q range .00825-.413 or .00621-.333 $Å^{-1}$. Total scattering plots were generated by summing the intensities over the entire q range of each data set and plotting against the frame number. Reduced data from contiguous or non-contiguous regions representing buffer alone or a combination of protein and buffer were analyzed by singular value decomposition and the protein scattering reconstructed by linear combination of the two most signification basis components with coefficients optimized by linearity of the Guinier region. Non-uniform errors were estimated by Savitzky-Golay smoothing using a 3^{rd} order polynomial with a window size of 11 points [222]. A constant was subtracted from the data to obtain a Porod Volume with theoretical molecular weight within 10% of the corresponding solvent-free protein. Kratky, Porod, and Guinier plots and fits were calculated within the in-house Mac OSX application DELA. In test cases using autoinhibited Grp1 as well as isolated PH and Arf6 constructs, corrected data produced scattering, Guinier R_g, Porod Volume, molecular weight, and D_{max} values nearly identical to theoretical values calculated from the crystal structure using CRYSOL.

Dummy Atom Modeling

Programs in the ATSAS packages were used for data modeling [213]. CRYSOL with constant subtraction was used to calculate theoretical scattering curves from crystal structures. PRIMUS [219] and maximum entropy method fitting in DELA were used for P(r) distribution calculation. Distributions were calculated using all of the data points, and D_{max} values were adjusted in Primus until distance frequencies naturally descended to 0. GNOM output files containing the P(r) distributions were used as input for DAMMIF and GASBOR [215, 216]. One hundred cycles of DAMMIF were performed. Averaging through the DAMAVER suite was performed in groups of ten, and these envelopes averaged and filtered. Average NSD values reported are the final values between the groups of ten. Within each group of ten structures, values ranged from 0.4 - 0.6 for autoinhibited Grp1 constructs, and 1.0-1.5 for Grp1-Arf6 tandems.

Ensemble Optimization Method and Rigid Body Modeling

The EOM suite integration of the programs RANCH and GAJOE were used to generate 10,000 structures based on crystallographic subunits and refine a best-fit ensemble. Representative individual structures had a large range of chi-values and were refined to SAXS data for the both full length and relevant truncated constructs using the program CORAL [242]. For both EOM and CORAL, the Grp1 PH domain and the Arf6 module were fixed, with residues representing the linker regions between subunits and N/C-terminal histidine tags. Final solutions in which the Grp1-Arf6 interaction was not maintained were rejected.

CHAPTER V: DISCUSSION AND FUTURE DIRECTIONS

A Unified Mechanism for Relief of Autoinhibition

The work presented in this thesis clearly defines the molecular mechanisms of autoinhibition relief in Cytohesin GEFs by Arf GTPases. Importantly, this mechanism is compatible with autoinhibition relief through phosphorylation of Cytohesins by Protein Kinase C (PKC) [133, 148] and helps explain Grp1 mutants outside the Sec7 or phosphoinositide binding pocket of the PH domain that ablate Arf6-dependent cell spreading [243]. Arf-GTP binding may be a more stable form of Grp1 stimulation than phosphorylation by sequestering the polybasic region and orienting it for membrane insertion, whereas phosphorylation simply disfavors the autoinhibited state.

Additionally, the requirement of phosphoinositide binding for efficient Arf-GTP recognition and participation of the unique hairpin insertion known to insert into membranes point to environmental regulation of activity and evolution of cytohesins for nucleotide exchange at membrane environments. As 2G Cytohesin splice variants show 30-fold selectivity for PIP_3 over 3G isoforms, the temporal or tissue-specific expression levels of each isoform may play a role in regulating activity. Similarly, ARNO and CYTH1 are predominantly expressed in the 3G forms, but undergo PKC phosphorylation, raising the possibility that phosphorylation is a compensatory or specialized mechanism for stimulation in these isoforms. As a whole, current evidence has defined general mechanisms for cytohesin activity related to common Sec7 domain function and membrane recruitment via PH domain binding to phosphoinositides, with specific regulation related to PIP_3 and or Arf-GTP levels, PKC expression, and splice

variation. A current focus of ongoing research in GTPase biology will be to analyze mechanisms of cytohesin autoinhibition relief in the context of the known dimeric species, as well as analyze stimulation by other GTPases. These insights could lead to more informed studies in the development of small molecule cell biological tools or therapeutics [128, 135, 244, 245], or used to study other autoregulated GEFs.

Exploring Autoinhibition Relief with Cytohesin Dimers

The heptad repeat (HR) region in Cytohesins interacts with scaffolding proteins such as CASP and GRASP as well as the associated protein Grsp1 [130-132]. In addition to heterodimers with these proteins, the Cytohesin family members ARNO and Grp1 participate in HR-HR homodimers at physiologically relevant concentrations. Moreover, recent evidence of a HR-PH domain autoinhibited dimer in ARNO centering around an Akt phosphorylation site at Thr 276 also implicates a physiologically relevant homodimer [134].

The role of these dimers in autoregulatory mechanisms is not completely known. Autoinhibition and recruitment to membrane environments is unaffected by homo or heterodimerization *in vitro* [130, 133], but it is unknown whether there is an effect of dimers *in vivo* or in the context of Arf6-binding. Although the crystallographic dimer would preclude relief of autoinhibition by Arf6 and is not thought to occur at physiological concentrations, both former HR-HR and HR-PH dimers could be compatible with Arf6 binding. The HR-PH domain dimer recently isolated in cells is thought to have an added inhibitory role on catalytic activity [134].

The structural organization of these species is unknown, Heterodimers with the coiled-coil domain of Grsp1 are believed to be antiparallel [130], but homodimers are structurally uncharacterized. The experiments described in this work led to the hypothesis that coiled-coil homodimers could have stable structural organization of the Sec7 domain, going from a membrane distal to a membrane proximal location upon stimulation. To structurally characterize HR-dependent Cytohesin dimers, initial SEC-SAXS was performed on dimeric forms of Grp1 and Arno in autoinhibited ($Grp1_{13-399}$, $Arno_{2-400}$) and active ($Grp1_{13-390}$) forms [133]. All constructs produced a single SEC UV absorbance and scattering peaks that could be analyzed by SVD-LC correction followed by Guinier and P(r) analysis (Figure V.1 and V.2). Data were compared to monomer and dimer data from experiments with autoinhibited $Grp1_{63-399}$ (Figure III.9). As might be expected by molecular weight rankings, $Grp1_{13-399}$ and $Arno_{2-400}$ had similar R_g values (53.8 and 52.6 Å) higher than $Grp1_{13-390}$ (48.6 Å) and the $Grp1_{63-399}$ (40.0 Å) high concentration dimer. $Grp1_{13-399}$ and $Grp1_{13-390}$ produced broad P(r) distributions pointing to a consistently elongated structure without large deviation despite the large increases in activity in this construct [133]. (Figure V.2C). Shape envelopes calculated for $Grp1_{13-399}$ and $Arno_{2-400}$ confirmed the appearance of elongated, planar average structures with at least partial symmetry.

Overall, these data show that dimeric Cytohesins form elongated architectures, perhaps with flexibility or unique folds within the dimer interface. These species are incompatible with more compact structures seen in crystallographic studies (Figure III.9), and point to either a parallel or anti-parallel coiled coil. More studies are needed to

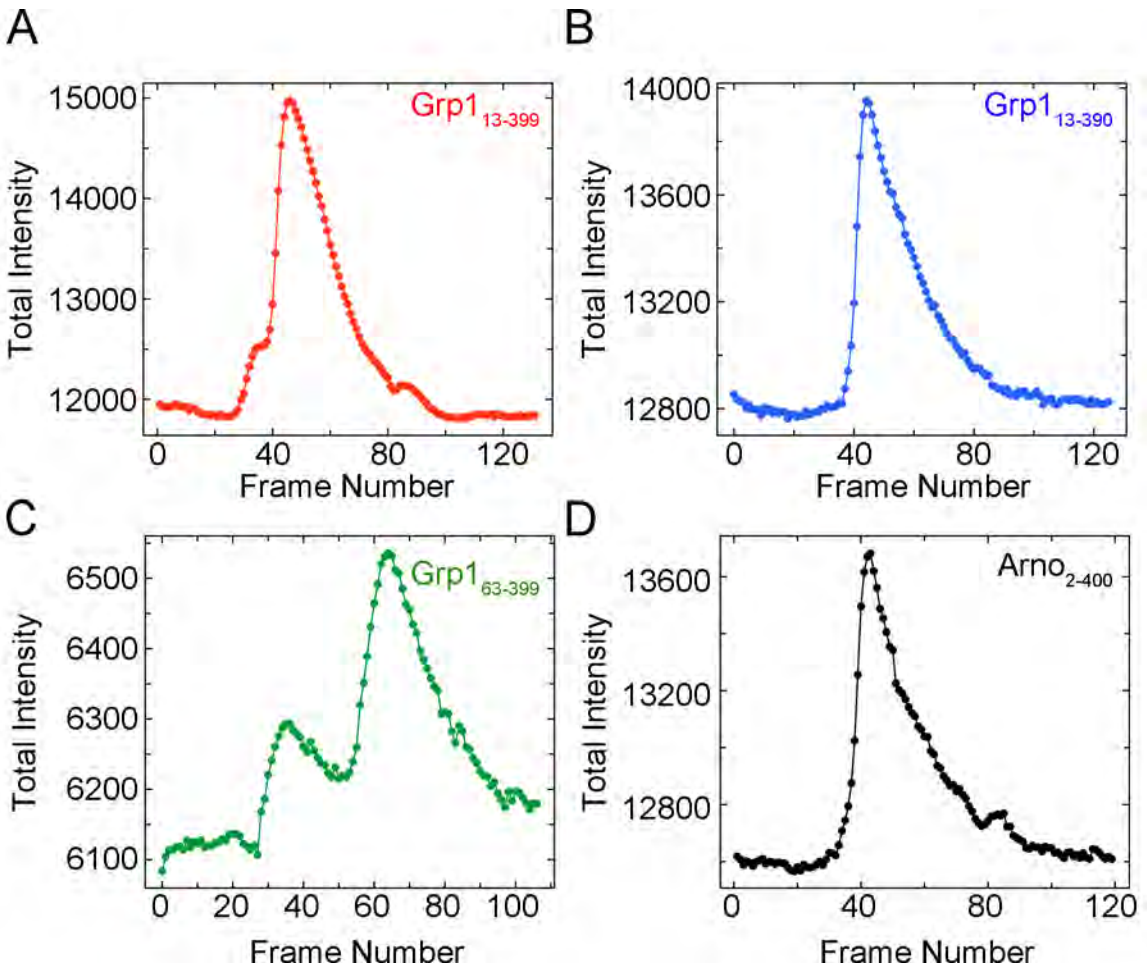

Figure V.1: SAXS Total Intensity Chromatograms for Dimeric Grp1 Constructs. Total Intensity Chromatograms for $Grp1_{13-399}$ (A), $Grp1_{13-390}$ (B), $Grp1_{63-399}$ (C), and $ARNO_{2-400}$ (D)

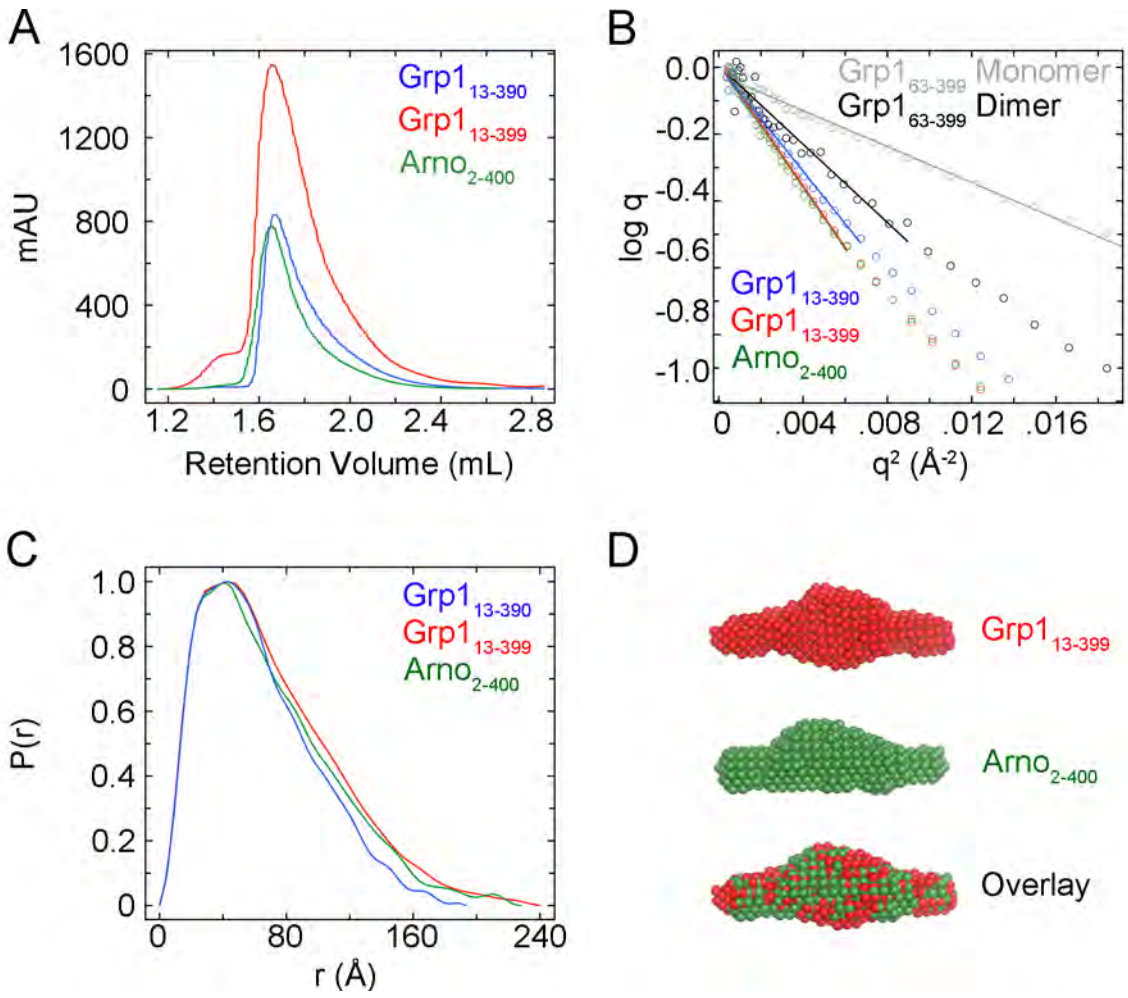

Figure V.2: Analysis of Grp1 Dimers by SEC-SAXS. A) UV Chromatograms from SEC-SAXS on constructs as indicated. B) Guinier analysis. C) P(r) Distributions of $Grp1_{13-399}$ and $Arno_{2-400}$. D) DAMMIF shape envelopes of $Grp1_{13-399}$ and $Arno_{2-400}$.

assess flexibility in these molecules and to assign the coiled coil architecture. Modeling analysis with hypothetical parallel or antiparallel coiled-coil models in an autoinhibited or active Grp1 molecule may help decipher this overall architecture. EOM analysis and biochemical analysis offers an interesting way to propose and test the feasibility of these models. Alternatively, crystallographic studies of the full length or HR regions of cytohesins may also aid in this project.

While homo- and heterodimers maintain monomeric levels of phosphoinositide binding and catalytic activity a potential role of these dimers in autoinhibition relief has yet to be probed. It is unknown whether Arf6 can relieve autoinhibition of these species, although given integration of relief of autoinhibition with membrane recruitment, the ability of Arf6 to recruit ARNO to the PM suggests that it does indeed function similarly in these cases. As the predicted coiled coil extends completely to the Sec7 domain, dimerization may stabilize the Sec7 domain after Arf6 binding (Figure V.3). Under full membrane association conditions, relief of autoinhibition by Arf6 would provide a rotation of the PH domain and stabilization of the Sec7 domain, rather than a large fluctuation of the Sec7 domain.

Mechanisms of Membrane Recruitment for Cytohesin Recruitment by Arf6

Arf6 can recruit Cytohesins to the plasma membrane upon PIP_3 upregulation by binding to the allosteric site [31, 75]. Although molecular determinants in membrane recruitment of Arfs and Cytohesins alone has been extensively studied [60, 61, 111, 124, 234, 246], additional factors in stabilization of the physiological complex of Cytohesins

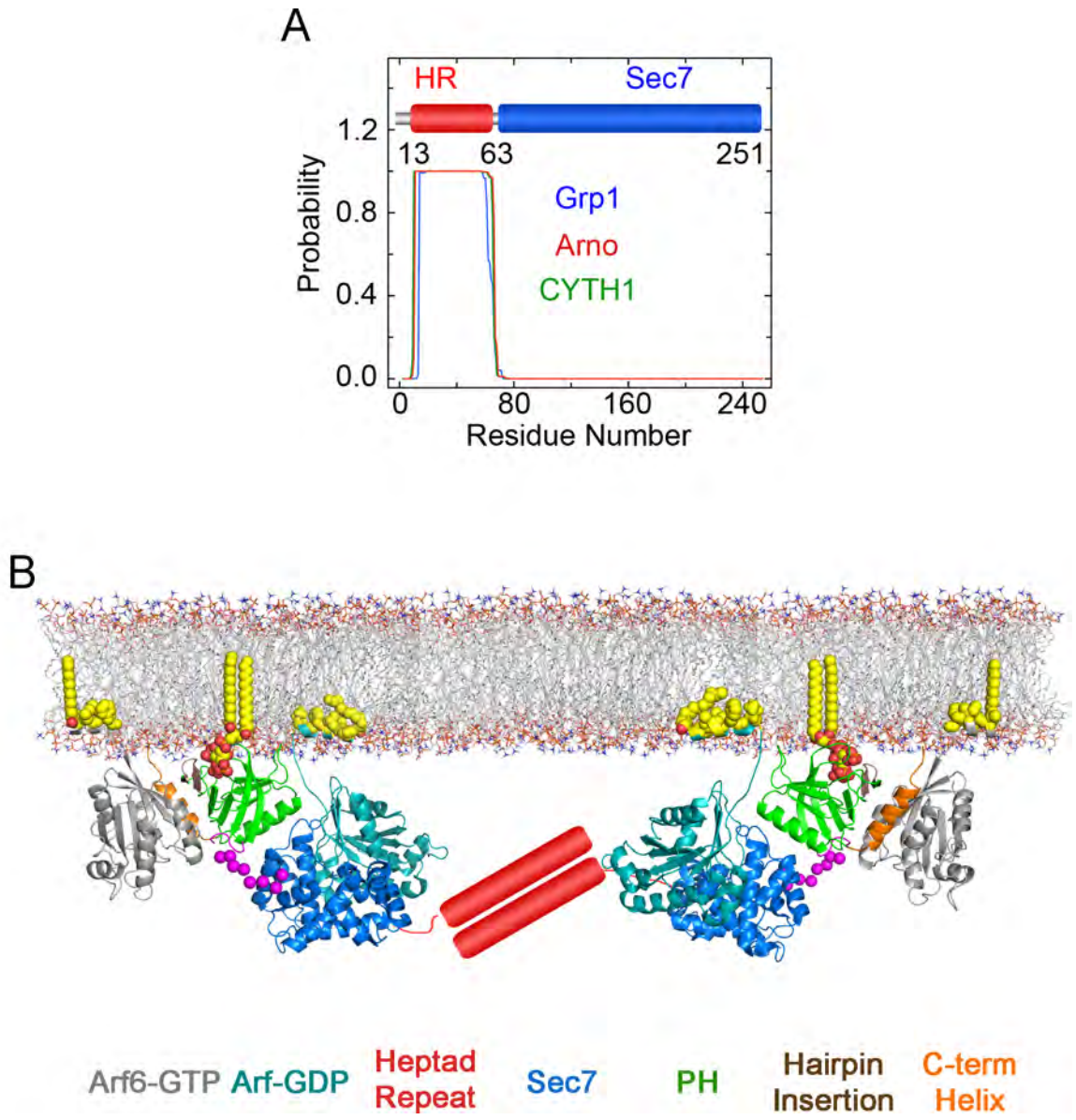

Figure V.3: Model for Autoinhibition Relief in Dimeric Cytohesins. A) Probability of coiled coil formation in Cytohesins using the prediction program COILS with a window size of 28. B) Model depicting stabilization of membrane proximal nucleotide exchange by Arf stimulated Cytohesins by substrate binding or HR region dimerization. Model was similar to Figure IV.9, but with manual manipulation of Sec7 domain. Arf6-GDP (PDB 1E0S) was docked into the Sec7 binding site via superposition with the structure of the ARNO Sec7 in complex with Arf1-GDP (PDB ID 1R8S) and the C-terminus oriented for membrane insertion in PyMol.

with active and substrate Arf at membranes have not been investigated. SAXS experiments presented in Chapter IV suggest that allosteric site binding also promotes a membrane proximal position of the Sec7 domain upon Arf6 binding. It remains to be seen what role rearrangement of the Sec7 domain and/or substrate binding plays in the overall mechanism integrating membrane targeting with autoregulation. Arf-GDP binding may help stabilize a membrane-proximal conformation, providing a concerted mechanism to create stable microenvironments of specialized GEF activity. Stabilization of Arf-GDP association through association of the Sec7-PH domain linker has been shown in the GEF BRAG2 [87]. In addition, Arf6-bound conformations of the Sec7 domain may itself contribute to membrane association through direct membrane insertion, such as a lysine patch at α1 and α2.

Mutational analyses of dimeric and monomeric Grp1 using liposome partitioning assays in the presence of myristoylated Arf6-GTP may provide insight into determinants in the Sec7 domain involved in membrane recruitment. Additionally, biochemical and SAXS experiments using Grp1-Arf6 tandems carrying the Grp1 E161K substitution and in the presence of Arf-GDP may help determine the stabilization of the Sec7 domain by substrate binding cooperatively or independently of the membrane environment. These and other biochemical and structural studies with dimeric and substrate-bound cytohesins will define the sequential and concerted mechanisms of relief of autoinhibition and membrane recruitment by GTP-bound Arfs and Arls.

Relief of Autoinhibition by Arl4

The Grp1-Arf6 crystal structure showed a large binding epitope on Arf6 that is conserved amongst broader Arf/Arl family members. Indeed, Arf1-GTP has been shown to relieve autoinhibition in the context of full length, myristoylated Arfs and liposome systems, albeit at reduced levels [151, 234]. Despite this, stimulation in the absence of membranes was severely reduced with class I and class II Arfs, and could not be rescued even with complete substitution of the Arf1 switch region with that of Arf6. This selectivity was 8-10 fold in experiments using soluble Arf effectors, but was almost 1000 fold for membrane associating Arf1 and Arf6 (Chapter II). One explanation for this selectivity is a simple electrostatic barrier to binding. Arf6 has a pI of 8.9, whereas other Arfs are significantly more acidic. Indeed, neutralization of electropositive charges in Grp1 by phosphoinositide binding may facilitate binding to only basic Arf molecules. By extension, this binding modality may represent a common mechanism of autoinhibition relief by Arfs.

One Arf protein hypothesized to act similarly to Arf6 is Arl4d, which has also been shown to bind to the Cytohesin PH domain and recruit ARNO to the PM in a PIP3 dependent manner [32, 75]. Recruitment stimulated Arf6-GDP activation at the PM, but this effect has not been separated from recruitment to the excluded volume of membrane environments. Arl4d is 40% identical to Arf6, considerably less than other class I and II Arfs, but has very high pI of 10.0, in part due to the presence of a C-terminal polybasic region important for recruitment. Arl4 may also have more robust membrane insertion due to a four-residue insertion in the ISW region compared to other Arfs [24]. This

would promote association of the N-terminus by disfavoring a buried conformation of the amphipathic helix. Crystal structures have shown similar overall GTP-bound structures in Arfs and Arls [71, 247, 248], and the Grp1 PH-Arf6 crystal structure is compatible with recruitment of Cytohesins to the PM via a C-terminal extension, which is also oriented toward a membrane binding surface of the complex.

Because of the highly basic nature of this protein and poor expression of soluble material, initial experiments have proven difficult; however, we have shown relief of autoinhibition in solution experiments using full length Arl4d and soluble substrate reporter with liposome environments containing PIP_3 (Figure V.4). More experiments using truncation of the N and/or C-termini as well as Grp1 mutants in the presence or absence of membrane components are expected to elucidate whether the mechanism of action is similar to Arf6. Additionally, as the requirement for phosphoinositide in this interaction is not clear [32, 75], *in vitro* experiments as designed in Chapter II with liposome compositions will help distinguish allosteric binding effects from membrane recruitment.

Mechanisms for Autoinhibition Relief in the Pathogenic GEF RalF

Many pathogenic bacteria secrete a myriad of effector proteins which help it avoid host defenses and hijack cellular machinery [7]. As such, understanding these pathways may help treat infections and offer important insight into general mechanisms for membrane trafficking [7]. One particular example of this is the Arf GEF RalF from *Legionella pneumophila* and *Rickettsia rickettsii*. Arf-GDP binding to a Sec7 domain is

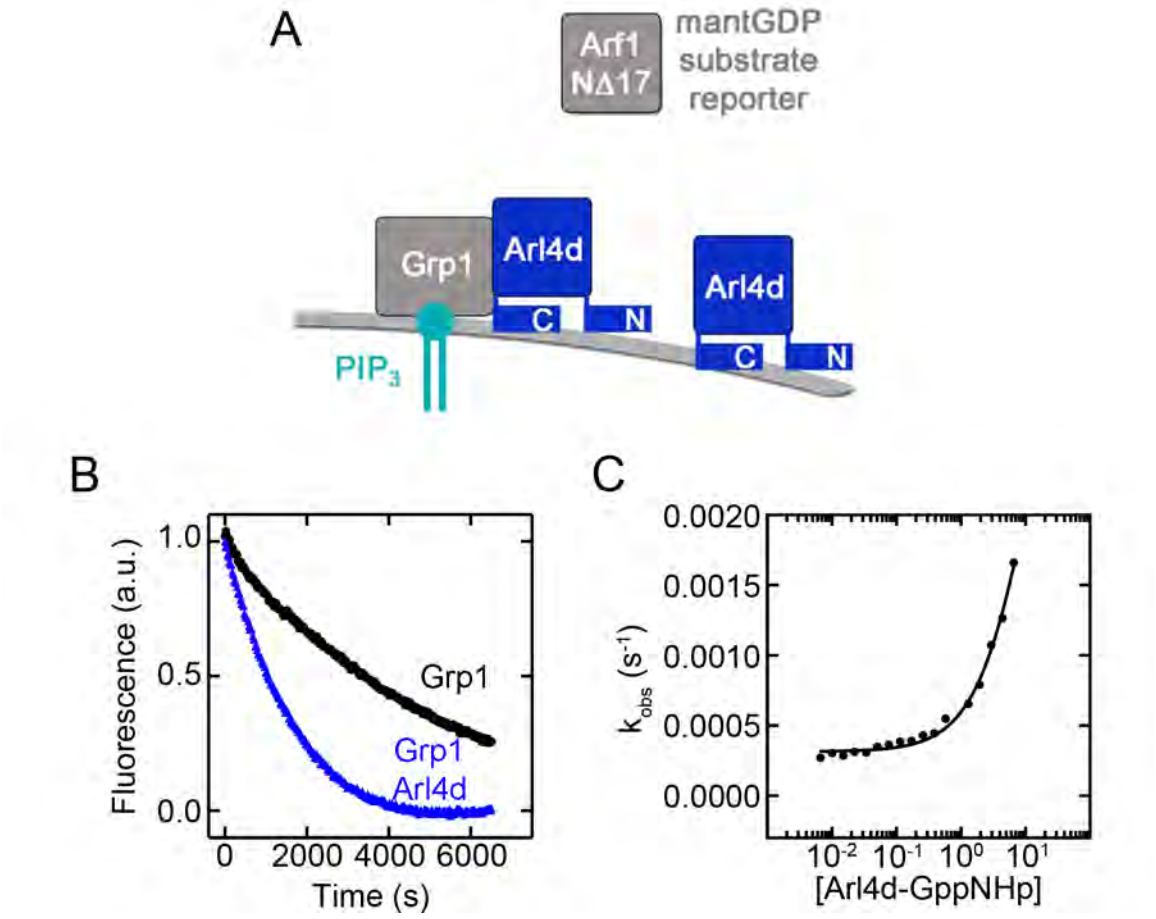

Figure V.4: Stimulation of Grp1 Activity by Arl4d-GppNHp. A) Schematic of kinetic experiments measuring mant-GDP dissociation from NΔ17Arf1 with Grp1 in the presence of PIP_3 containing liposomes and Arl4d-GppNHp. B) Representative dissociation of mant-GDP from 1 μM NΔ17Arf1 with 125 nM $Grp1_{63-399}$ in the absence or presence of 6.25 μM Arl4d-GppNHp and .5 mM 97% POPC: 3% PIP_3 liposomes C) Dependence of nucleotide release on concentration of Arl4d.

blocked by a C-terminal capping domain that binds phosphoinositides [26]. Membrane environments have been shown to dramatically increase nucleotide exchange, but as activity assays have all included full length myristoylated substrates, autoinhibition per se has never definitively been measured *in vitro*. Because the potential phosphoinositide binding site based on homology modeling is far away from the substrate binding site, it is unclear how or if phosphoinositide binding may relieve autoinhibition. Additionally, while *Legionella* RalF shows a binding preference for PI4P, it is unknown whether other phosphoinositides can contribute to this mechanism, and phosphoinositide binding for *Rickettsia* RalF may be promiscuous among negatively charged lipids. Key residue differences in an 'aromatic cluster' have been shown to underlie these differences [249].

In an effort to clarify these relationships, preliminary liposome sedimentation profiling was performed with either *Rickettsia rickettsii* or *Legionella pneumophila* protein incorporating WT amino acid sequences or reciprocal mutations to convert aromatic cluster regions (Figure V.5). Rickettsia normally has lysines at these positions (3K), while Legionella has Phe, Gln, and Asn (FQN) at these sites. Wild type FQN *Legionella* RalF showed slight preference for PI4P, but significant association with other phosphoinositides at higher concentrations. In contrast, the wild type 3K *Rickettsia* RalF partitioned most strongly with PIP_2 and PIP_3, but also had notable cosedimentation with PI4P. All 3K variants had increased association with phosphoinositide-containing liposomes, but importantly did not override the PI4P preference at low concentrations. These assays clearly define PI4P preference in *Legionella* RalF and a lipid charge sensing mechanism for *Rickettsia* RalF in a liposome environment, consistent with previous data

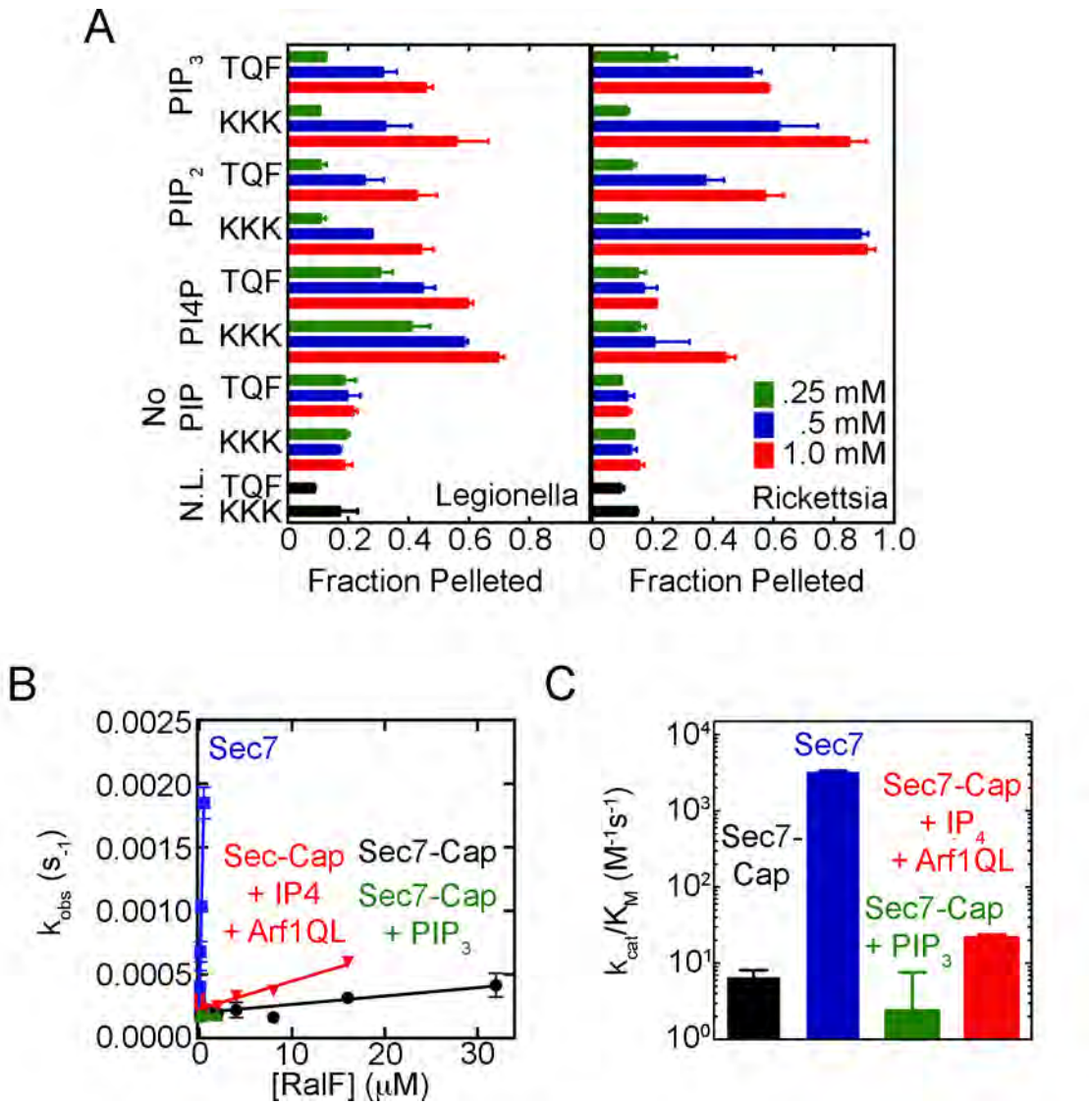

Figure V.5: Liposome Partitioning and Catalytic Activity of RalF. A) Sedimentation assays with varying concentrations of liposomes containing 97% POPC and 3% PI4P, PIP_2, or PIP_3 and Legionella or Rickettsia RalF with the indicated aromatic cluster residues. B) Dependence of kinetic rate of mant-GDP release from NΔ17Arf1 on the concentration of *Rickettsia* RalF Sec7 or Sec7-Cap constructs in the absence of presence of either .5 mM PIP_3-containing liposomes or 10 μM IP_4 and 80 μM NΔ17Arf1(Q71L). C) Catalytic efficiencies from the experiments in B.

using lipid overlay assays [249]. Addtional biochemical and structural assays will further define these mechanistic differences at the molecular level.

Preliminary assays were also performed to test autoinhibition relief mechanisms in RalF. To separate mechanisms of membrane recruitment from allosteric activation, soluble Arf1ΔN17 was used as substrate for *Legionella* RalF (Figure V.5B and C). Experiments using the Sec7 and Capping domain (Sec7-Cap) showed very low activity that was almost undetectable at moderate concentrations. Conversely, deletion of the capping domain showed almost 1000 fold increases in activity. Liposomes containing PIP_3 at saturating levels based on the sedimentation assays (0.5 mM liposomes with up to 1 μM protein) did not increase activity, pointing to an external factor required for full relief of autoinhibition. Interesting candidates include Arfs, which are substrates for RalF or Rab1, which is a common target of many *Legionella* effectors. In support of this, the head group for PIP_3, IP_4 in combination with truncated Arf1Q71L, which can not associate with membranes, did have a modest improvement in catalytic efficiency (Figure V.5C). More biochemical studies screening effector GTPases in soluble and membrane-bound forms and structural studies of the effector with RalF will shed light into the molecular mechanisms of autoinhibition relieve in this system.

Overall, the unique PI4P preference in *Legionella* RalF and another effector DrrA [250] as well as robust autoinhibition greater than that seen in any mammalian GEF system highlight mechanisms that work together to efficiently hijack the ER machinery rich in PI4P. *Rickettsia* RalF, while having similar autoregulatory properties, has evolved a charge sensing C-terminal domain for perhaps more general or widespread membrane

recruitment and autoregulation. It will be exciting to learn potential mechanisms of relief of RalF autoinhibition, potentially incorporating host GTPases like Arf1 that are upregulated at ER-like membranes necessary for the *Legionella* life cycle.

Overall Implications and Concluding Remarks

As mentioned, autoregulation of GEFs is a common theme in GTPase biology, seen in GEFs for a variety of Ras superfamily GTPases [87, 133, 137, 139, 233]. These GEFs are emerging as key players or therapeutic targets for such diseases as colorectal cancer, diabetes, or bacterial infection [47, 244, 245, 251, 252]. Modes of autoinhibition relief include phosphorylation by serine/threonine or tyrosine kinases, lipid binding, and effector binding. While phosphorylation recognition is very specific, phosphoinositide and effector binding may offer a more fine tuned gradient of activity, controlled as much by local environment as specificity of a particular interaction. Additionally the integration of membrane recruitment with mechanisms for relief of autoinhibition represents a unique form of regulation beyond simple coincidence detection seen in cellular trafficking. A full understanding of the intricate regulation of these proteins will lead to more targeted therapies and a better understanding of the complex processes involved in cellular transport.

Materials and Methods

Cloning and Protein Purification

DNA constructs were amplified using PCR with Vent polymerase (NEB), digested with BamH1, and SalI/XhoI and ligated into a modified pET15b vector. Mutants were generated using whole plasmid PCR with QuikSolution (Stratagene) and DpnI (NEB) digestion. After transformation of BL21(DE3) cells (Novagen) with plasmids, cells were grown in 2×YT with 100 mg/L ampicillin to an OD_{600} of 0.2-0.6 or 1.0-1.2 and expression induced with 50 μM or 1mM IPTG at 18°C for 16hrs or 37°C for 3 hrs. After harvest, cells were resuspended in buffer (50mM Tris pH 8.5 or 8.8, 150mM NaCl, 2mM $MgCl_2$, 0.05% 2-mercaptoethanol) and frozen at -80°C. Cells were thawed and incubated with 0.1 mM PMSF, 0.2 mg/ml lysozyme, and 0.01 mg/ml protease free DNAse I (Worthington). After sonication, lysates were supplemented with 0.5% Triton X-100 and centrifuged at 30,000 × g for 1 hr. Proteins were purified on batch Ni-NTA beads followed by ion exchange on HiTrap Q or SP and gel filtration on Superdex-75 or 200 (GE Healthcare).

RalF constructs consisting of C-terminal Capping Domain and in the pMal5c vector were a gift from the lab of Dr. Craig Roy. BL21(DE3) cells were transformed and induced as above. Cells were lysed in buffer (20 mM Hepes pH 7.4, 300 mM NaCl, 1 mM $MgCl_2$, 0.05% 2-mercaptoethanol) and purified on maltose binding protein, Q or SP, and gel filtration on Superdex-200 (GE Healthcare). Capping domain constructs purified as soluble aggregates and full length proteins as monomeric species, as measured by analytical gel filtration.

SEC-SAXS

SEC-SAXS was performed at Bio-CAT, Sector 18-ID at Advanced Photon Source (APS) in Argonne National Labs, Argonne, IL. 100 – 500 ul (10-20 mg/ml) or 100ul (10 mg/ml) samples were loaded onto 24 or 3 ml Superdex-200 columns size exclusion columns, respectively (GE Healthcare). Columns were connected in line with the beam and 1 s exposures were taken at 5 s intervals during elution. Data was monitored by UV chromatogram and showed consistent and clear separation of monomeric protein on both columns. Samples were calculated to be between 1-3 mg/ml at the peak of the protein elution profile.

Data Processing

A detailed description of data processing including method development will be given in Appendix A. Briefly, data images were radially averaged over the q range .00825-.413 or .00621-.333. Data normalization, and buffer correction by singular value decomposition and linear combination, and constant background subtraction were performed as described ([218] and Chapter III)

Dummy Atom Modeling

Programs in the ATSAS packages were used for data modeling [213]. Crysol was used to calculate theoretical scattering curves from crystal structures. PRIMUS [219] and maximum entropy method fitting were used for P(r) distribution calculation, and the

GNOM result used as input for DAMMIF and GASBOR [215, 216]. One hundred cycles of DAMMIF were performed. Averaging through the DAMAVER suite was performed in groups of ten, and these envelopes averaged and filtered.

Liposome Preparation

Liposomes containing 67-70% palmitoyloleolylphosphatidylcholine (Avanti), 30% palmitoyloleolylphosphatidylserine (Avanti), and 3% PI4P, PIP_2, or PIP_3 (Cell Signals) were prepared as follows. Lipids dissolved in 2:1 chloroform:methanol were mixed and dried under an argon stream. For GEF assays, lipid mixtures were rehydrated in 20 mM Tris HCl, pH 8.0, 150 mM NaCl, and 2 mM $MgCl_2$. Large Unilamelar Vesicles (LUVs) were prepared by extrusion through 0.1 μm filters (Avanti) after 10 freeze-thaw cycles in liquid nitrogen. For GEF Assays, the liposomes were resuspended in GEF Assay Buffer (see below). For liposome partitioning, lipid mixtures were rehydrated in 20 mM Hepes, pH 7.5, and 170 mM sucroseFollowing 10 freeze-thaw cycles and extrusion as above, the liposomes were diluted 4:1 with 20 mM Hepes, pH 7.5, and 100 mM KCl, pelleted by ultracentrifugation at 100,000×g for 30 minutes, and resuspended at a total lipid concentration of 5 mM.

Liposome Partitioning Assays

3 μM RalF was incubated with 0, .25, .5, or 1 mM sucrose loaded liposomes in 100 reactions in buffer (20 mM Hepes, pH 7.5, and 100 mM KCl) and incubated for 30 minutes. Reactions were centrifuged at 100,000 x *g* and supernatants and pellets

analyzed by SDS-PAGE followed by Coomassie Blue staining and densitometry quantification in ImageJ.

Nucleotide Loading

Arl4d was incubated overnight at 4°C in 20 mM Tris, pH 8.0, 150 mM NaCl, 5 mM EDTA, and 1 mM DTT with a 10 fold molar excess of GppNHp. $MgCl_2$ was added to a final concentration of 10 mM and excess nucleotide removed by gel filtration on Superdex-75 (GE Healthcare) or D-Salt column (Pierce ThermoFisher).

GEF Assays

The fluorescence decrease upon mantGDP dissociation from Arf1NΔ17 was monitored in 20 mM Tris, pH 8.0, 150 mM NaCl, 2 mM $MgCl_2$ and 250 μM or 1 mM GppNHp. Grp1 or *Ricketsia* RalF with or without active Arf/Arl GTPases and/or head groups/liposomes was formatted into 96 well half area microplates (Corning) and incubated for 16-24 hrs at 25°C. Reactions were initiated by addition of 1 μM Arf1NΔ17-mantGDP and monitored using a Safire microplate spectrophotometer (Tecan) with excitation and emission wavelengths of 360 nm and 440 nm, respectively. Observed rates constants (k_{obs}) were obtained by fitting with $I_t = (I_0 - I_\infty) \exp(-k_{obs}\, t) + I_\infty$, where I_t, I_0, and I_∞ are the emission intensities at times t, t = 0, and as $t \rightarrow \infty$. Catalytic efficiencies (k_{cat}/K_M) were obtained from the slope of a linear least squares fit with $k_{obs} = (k_{cat}/K_M)[GEF] + k_{intr}$, where k_{intr} is the intrinsic dissociation rate constant.

Works Cited

1. Bierings, R., M. Beato, and M.J. Edel, *An endothelial cell genetic screen identifies the GTPase Rem2 as a suppressor of p19ARF expression that promotes endothelial cell proliferation and angiogenesis.* J Biol Chem, 2008. **283**(7): p. 4408-16.
2. Fan, Y., et al., *Mutations in a member of the Ras superfamily of small GTP-binding proteins causes Bardet-Biedl syndrome.* Nat Genet, 2004. **36**(9): p. 989-93.
3. Uchiyama, K., H. Miyata, and S. Sakaguchi, *Disturbed vesicular trafficking of membrane proteins in prion disease.* Prion, 2013. **7**(6).
4. Dowling, J.J., E.M. Gibbs, and E.L. Feldman, *Membrane traffic and muscle: lessons from human disease.* Traffic, 2008. **9**(7): p. 1035-43.
5. Accili, D., et al., *A mutation in the insulin receptor gene that impairs transport of the receptor to the plasma membrane and causes insulin-resistant diabetes.* EMBO J, 1989. **8**(9): p. 2509-17.
6. Aridor, M. and L.A. Hannan, *Traffic jam: a compendium of human diseases that affect intracellular transport processes.* Traffic, 2000. **1**(11): p. 836-51.
7. Kahn, R.A., H. Fu, and C.R. Roy, *Cellular hijacking: a common strategy for microbial infection.* Trends Biochem Sci, 2002. **27**(6): p. 308-14.
8. Ge, J. and F. Shao, *Manipulation of host vesicular trafficking and innate immune defence by Legionella Dot/Icm effectors.* Cell Microbiol, 2011. **13**(12): p. 1870-80.
9. Stenmark, H., *Rab GTPases as coordinators of vesicle traffic.* Nat Rev Mol Cell Biol, 2009. **10**(8): p. 513-25.
10. Donaldson, J.G. and C.L. Jackson, *ARF family G proteins and their regulators: roles in membrane transport, development and disease.* Nat Rev Mol Cell Biol, 2011. **12**(6): p. 362-75.
11. Cherfils, J. and M. Zeghouf, *Regulation of small GTPases by GEFs, GAPs, and GDIs.* Physiol Rev, 2013. **93**(1): p. 269-309.
12. Mizuno-Yamasaki, E., F. Rivera-Molina, and P. Novick, *GTPase networks in membrane traffic.* Annu Rev Biochem, 2012. **81**: p. 637-59.
13. Grosshans, B.L., D. Ortiz, and P. Novick, *Rabs and their effectors: achieving specificity in membrane traffic.* Proc Natl Acad Sci U S A, 2006. **103**(32): p. 11821-7.
14. Resh, M.D., *Trafficking and signaling by fatty-acylated and prenylated proteins.* Nat Chem Biol, 2006. **2**(11): p. 584-90.
15. Alland, L., et al., *Dual myristylation and palmitylation of Src family member p59fyn affects subcellular localization.* J Biol Chem, 1994. **269**(24): p. 16701-5.
16. Resh, M.D., *Myristylation and palmitylation of Src family members: the fats of the matter.* Cell, 1994. **76**(3): p. 411-3.

17. Sigal, C.T., et al., *Amino-terminal basic residues of Src mediate membrane binding through electrostatic interaction with acidic phospholipids.* Proc Natl Acad Sci U S A, 1994. **91**(25): p. 12253-7.
18. Buser, C.A., et al., *Membrane binding of myristylated peptides corresponding to the NH2 terminus of Src.* Biochemistry, 1994. **33**(44): p. 13093-101.
19. Murray, D., et al., *Electrostatics and the membrane association of Src: theory and experiment.* Biochemistry, 1998. **37**(8): p. 2145-59.
20. Hancock, J.F., H. Paterson, and C.J. Marshall, *A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane.* Cell, 1990. **63**(1): p. 133-9.
21. Schroeder, H., et al., *S-Acylation and plasma membrane targeting of the farnesylated carboxyl-terminal peptide of N-ras in mammalian fibroblasts.* Biochemistry, 1997. **36**(42): p. 13102-9.
22. Blumer, J., et al., *Specific localization of Rabs at intracellular membranes.* Biochem Soc Trans, 2012. **40**(6): p. 1421-5.
23. Antonny, B., et al., *N-terminal hydrophobic residues of the G-protein ADP-ribosylation factor-1 insert into membrane phospholipids upon GDP to GTP exchange.* Biochemistry, 1997. **36**(15): p. 4675-84.
24. Pasqualato, S., et al., *The structural GDP/GTP cycle of human Arf6.* EMBO Rep, 2001. **2**(3): p. 234-8.
25. Walch-Solimena, C., R.N. Collins, and P.J. Novick, *Sec2p mediates nucleotide exchange on Sec4p and is involved in polarized delivery of post-Golgi vesicles.* J Cell Biol, 1997. **137**(7): p. 1495-509.
26. Amor, J.C., et al., *The structure of RalF, an ADP-ribosylation factor guanine nucleotide exchange factor from Legionella pneumophila, reveals the presence of a cap over the active site.* J Biol Chem, 2005. **280**(2): p. 1392-400.
27. Poteryaev, D., et al., *Identification of the switch in early-to-late endosome transition.* Cell, 2010. **141**(3): p. 497-508.
28. Zhu, H., Z. Liang, and G. Li, *Rabex-5 is a Rab22 effector and mediates a Rab22-Rab5 signaling cascade in endocytosis.* Mol Biol Cell, 2009. **20**(22): p. 4720-9.
29. Stalder, D. and B. Antonny, *Arf GTPase regulation through cascade mechanisms and positive feedback loops.* FEBS Lett, 2013. **587**(13): p. 2028-35.
30. Margarit, S.M., et al., *Structural evidence for feedback activation by Ras.GTP of the Ras-specific nucleotide exchange factor SOS.* Cell, 2003. **112**(5): p. 685-95.
31. Cohen, L.A., et al., *Active Arf6 recruits ARNO/cytohesin GEFs to the PM by binding their PH domains.* Mol Biol Cell, 2007. **18**(6): p. 2244-53.
32. Li, C.C., et al., *ARL4D recruits cytohesin-2/ARNO to modulate actin remodeling.* Mol Biol Cell, 2007. **18**(11): p. 4420-37.
33. Franco, M., et al., *EFA6, a sec7 domain-containing exchange factor for ARF6, coordinates membrane recycling and actin cytoskeleton organization.* EMBO J, 1999. **18**(6): p. 1480-91.
34. Boshans, R.L., et al., *ADP-ribosylation factor 6 regulates actin cytoskeleton remodeling in coordination with Rac1 and RhoA.* Mol Cell Biol, 2000. **20**(10): p. 3685-94.

35. Koo, T.H., B.A. Eipper, and J.G. Donaldson, *Arf6 recruits the Rac GEF Kalirin to the plasma membrane facilitating Rac activation.* BMC Cell Biol, 2007. **8**: p. 29.
36. Allaire, P.D., et al., *Interplay between Rab35 and Arf6 controls cargo recycling to coordinate cell adhesion and migration.* J Cell Sci, 2013. **126**(Pt 3): p. 722-31.
37. Kobayashi, H. and M. Fukuda, *Rab35 regulates Arf6 activity through centaurin-beta2 (ACAP2) during neurite outgrowth.* J Cell Sci, 2012. **125**(Pt 9): p. 2235-43.
38. D'Souza-Schorey, C. and P. Chavrier, *ARF proteins: roles in membrane traffic and beyond.* Nat Rev Mol Cell Biol, 2006. **7**(5): p. 347-58.
39. Donaldson, J.G. and A. Honda, *Localization and function of Arf family GTPases.* Biochem Soc Trans, 2005. **33**(Pt 4): p. 639-42.
40. Kahn, R.A., et al., *Arf family GTPases: roles in membrane traffic and microtubule dynamics.* Biochem Soc Trans, 2005. **33**(Pt 6): p. 1269-72.
41. Volpicelli-Daley, L.A., et al., *Isoform-selective effects of the depletion of ADP-ribosylation factors 1-5 on membrane traffic.* Mol Biol Cell, 2005. **16**(10): p. 4495-508.
42. Beck, R., et al., *The COPI system: molecular mechanisms and function.* FEBS Lett, 2009. **583**(17): p. 2701-9.
43. Bonifacino, J.S. and J. Lippincott-Schwartz, *Coat proteins: shaping membrane transport.* Nat Rev Mol Cell Biol, 2003. **4**(5): p. 409-14.
44. De Matteis, M.A. and A. Godi, *Protein-lipid interactions in membrane trafficking at the Golgi complex.* Biochim Biophys Acta, 2004. **1666**(1-2): p. 264-74.
45. Mazelova, J., et al., *Ciliary targeting motif VxPx directs assembly of a trafficking module through Arf4.* EMBO J, 2009. **28**(3): p. 183-92.
46. Deretic, D., et al., *Rhodopsin C terminus, the site of mutations causing retinal disease, regulates trafficking by binding to ADP-ribosylation factor 4 (ARF4).* Proc Natl Acad Sci U S A, 2005. **102**(9): p. 3301-6.
47. Muralidharan-Chari, V., et al., *ARF6-regulated shedding of tumor cell-derived plasma membrane microvesicles.* Curr Biol, 2009. **19**(22): p. 1875-85.
48. Schafer, D.A., C. D'Souza-Schorey, and J.A. Cooper, *Actin assembly at membranes controlled by ARF6.* Traffic, 2000. **1**(11): p. 892-903.
49. Hashimoto, S., et al., *Requirement for Arf6 in breast cancer invasive activities.* Proc Natl Acad Sci U S A, 2004. **101**(17): p. 6647-52.
50. Fielding, A.B., et al., *Rab11-FIP3 and FIP4 interact with Arf6 and the exocyst to control membrane traffic in cytokinesis.* EMBO J, 2005. **24**(19): p. 3389-99.
51. Brown, F.D., et al., *Phosphatidylinositol 4,5-bisphosphate and Arf6-regulated membrane traffic.* J Cell Biol, 2001. **154**(5): p. 1007-17.
52. D'Souza-Schorey, C., et al., *A regulatory role for ARF6 in receptor-mediated endocytosis.* Science, 1995. **267**(5201): p. 1175-8.
53. Hiroyama, M. and J.H. Exton, *Localization and regulation of phospholipase D2 by ARF6.* J Cell Biochem, 2005. **95**(1): p. 149-64.
54. Lawrence, J.T. and M.J. Birnbaum, *ADP-ribosylation factor 6 regulates insulin secretion through plasma membrane phosphatidylinositol 4,5-bisphosphate.* Proc Natl Acad Sci U S A, 2003. **100**(23): p. 13320-5.

55. Wennerberg, K., K.L. Rossman, and C.J. Der, *The Ras superfamily at a glance.* J Cell Sci, 2005. **118**(Pt 5): p. 843-6.
56. Sievers, F. and D.G. Higgins, *Clustal Omega, accurate alignment of very large numbers of sequences.* Methods Mol Biol, 2014. **1079**: p. 105-16.
57. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Mol Syst Biol, 2011. **7**: p. 539.
58. Vetter, I.R. and A. Wittinghofer, *The guanine nucleotide-binding switch in three dimensions.* Science, 2001. **294**(5545): p. 1299-304.
59. Renault, L., et al., *Mechanism of domain closure of Sec7 domains and role in BFA sensitivity.* Biochemistry, 2002. **41**(11): p. 3605-12.
60. Liu, Y., R.A. Kahn, and J.H. Prestegard, *Structure and membrane interaction of myristoylated ARF1.* Structure, 2009. **17**(1): p. 79-87.
61. Liu, Y., R.A. Kahn, and J.H. Prestegard, *Dynamic structure of membrane-anchored Arf*GTP.* Nat Struct Mol Biol, 2010. **17**(7): p. 876-81.
62. Chavrier, P. and J. Menetrey, *Toward a structural understanding of arf family:effector specificity.* Structure, 2010. **18**(12): p. 1552-8.
63. Khan, A.R. and J. Menetrey, *Structural biology of Arf and Rab GTPases' effector recruitment and specificity.* Structure, 2013. **21**(8): p. 1284-97.
64. Schweitzer, J.K. and C. D'Souza-Schorey, *A requirement for ARF6 during the completion of cytokinesis.* Exp Cell Res, 2005. **311**(1): p. 74-83.
65. Isabet, T., et al., *The structural basis of Arf effector specificity: the crystal structure of ARF6 in a complex with JIP4.* EMBO J, 2009. **28**(18): p. 2835-45.
66. Zhang, T., et al., *Crystal structure of the ARL2-GTP-BART complex reveals a novel recognition and binding mode of small GTPase with effector.* Structure, 2009. **17**(4): p. 602-10.
67. Kahn, R.A. and A.G. Gilman, *The protein cofactor necessary for ADP-ribosylation of Gs by cholera toxin is itself a GTP binding protein.* J Biol Chem, 1986. **261**(17): p. 7906-11.
68. O'Neal, C.J., et al., *Structural basis for the activation of cholera toxin by human ARF6-GTP.* Science, 2005. **309**(5737): p. 1093-6.
69. Dell'Angelica, E.C., et al., *Molecular characterization of the protein encoded by the Hermansky-Pudlak syndrome type 1 gene.* J Biol Chem, 2000. **275**(2): p. 1300-6.
70. Bonifacino, J.S. and B.S. Glick, *The mechanisms of vesicle budding and fusion.* Cell, 2004. **116**(2): p. 153-66.
71. Hanzal-Bayer, M., et al., *The complex of Arl2-GTP and PDE delta: from structure to function.* EMBO J, 2002. **21**(9): p. 2095-106.
72. Hanzal-Bayer, M., M. Linari, and A. Wittinghofer, *Properties of the interaction of Arf-like protein 2 with PDEdelta.* J Mol Biol, 2005. **350**(5): p. 1074-82.
73. He, J., et al., *Molecular basis of phosphatidylinositol 4-phosphate and ARF1 GTPase recognition by the FAPP1 pleckstrin homology (PH) domain.* J Biol Chem, 2011. **286**(21): p. 18650-7.
74. Liu, Y., R.A. Kahn, and J.H. Prestegard, *Interaction of Fapp1 with Arf1 and PI4P at a Membrane Surface: An Example of Coincidence Detection.* Structure, 2014.

75. Hofmann, I., et al., *The Arl4 family of small G proteins can recruit the cytohesin Arf6 exchange factors to the plasma membrane.* Curr Biol, 2007. **17**(8): p. 711-6.
76. Greasley, S.E., et al., *The structure of rat ADP-ribosylation factor-1 (ARF-1) complexed to GDP determined from two different crystal forms.* Nat Struct Biol, 1995. **2**(9): p. 797-806.
77. Macia, E., et al., *The GDP-bound form of Arf6 is located at the plasma membrane.* J Cell Sci, 2004. **117**(Pt 11): p. 2389-98.
78. Casanova, J.E., *Regulation of Arf activation: the Sec7 family of guanine nucleotide exchange factors.* Traffic, 2007. **8**(11): p. 1476-85.
79. Goldberg, J., *Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching.* Cell, 1998. **95**(2): p. 237-48.
80. Achstetter, T., et al., *SEC7 encodes an unusual, high molecular weight protein required for membrane traffic from the yeast Golgi apparatus.* J Biol Chem, 1988. **263**(24): p. 11711-7.
81. Chardin, P., et al., *A human exchange factor for ARF contains Sec7- and pleckstrin-homology domains.* Nature, 1996. **384**(6608): p. 481-4.
82. Klarlund, J.K. and M.P. Czech, *Isolation and properties of GRP1, an ADP-ribosylation factor (ARF)-guanine nucleotide exchange protein regulated by phosphatidylinositol 3,4,5-trisphosphate.* Methods Enzymol, 2001. **329**: p. 279-89.
83. Betz, S.F., et al., *Solution structure of the cytohesin-1 (B2-1) Sec7 domain and its interaction with the GTPase ADP ribosylation factor 1.* Proc Natl Acad Sci U S A, 1998. **95**(14): p. 7909-14.
84. Cherfils, J., et al., *Structure of the Sec7 domain of the Arf exchange factor ARNO.* Nature, 1998. **392**(6671): p. 101-5.
85. Mossessova, E., J.M. Gulbis, and J. Goldberg, *Structure of the guanine nucleotide exchange factor Sec7 domain of human arno and analysis of the interaction with ARF GTPase.* Cell, 1998. **92**(3): p. 415-23.
86. Mossessova, E., R.A. Corpina, and J. Goldberg, *Crystal structure of ARF1*Sec7 complexed with Brefeldin A and its implications for the guanine nucleotide exchange mechanism.* Mol Cell, 2003. **12**(6): p. 1403-11.
87. Aizel, K., et al., *Integrated conformational and lipid-sensing regulation of endosomal ArfGEF BRAG2.* PLoS Biol, 2013. **11**(9): p. e1001652.
88. Rouhana, J., et al., *Kinetics of interaction between ADP-ribosylation factor-1 (Arf1) and the Sec7 domain of Arno guanine nucleotide exchange factor, modulation by allosteric factors, and the uncompetitive inhibitor brefeldin A.* J Biol Chem, 2013. **288**(7): p. 4659-72.
89. Renault, L., B. Guibert, and J. Cherfils, *Structural snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor.* Nature, 2003. **426**(6966): p. 525-30.
90. Kahn, R.A., et al., *Consensus nomenclature for the human ArfGAP domain-containing proteins.* J Cell Biol, 2008. **182**(6): p. 1039-44.

91. Nie, Z. and P.A. Randazzo, *Arf GAPs and membrane traffic.* J Cell Sci, 2006. **119**(Pt 7): p. 1203-11.
92. Ismail, S.A., et al., *The structure of an Arf-ArfGAP complex reveals a Ca2+ regulatory mechanism.* Cell, 2010. **141**(5): p. 812-21.
93. Cook, W.J., O. Senkovich, and D. Chattopadhyay, *Structure of the catalytic domain of Plasmodium falciparum ARF GTPase-activating protein (ARFGAP).* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2011. **67**(Pt 11): p. 1339-44.
94. Bai, M., et al., *Mechanistic insights into regulated cargo binding by ACAP1 protein.* J Biol Chem, 2012. **287**(34): p. 28675-85.
95. Goldberg, J., *Structural and functional analysis of the ARF1-ARFGAP complex reveals a role for coatomer in GTP hydrolysis.* Cell, 1999. **96**(6): p. 893-902.
96. Itoh, T. and T. Takenawa, *Mechanisms of membrane deformation by lipid-binding domains.* Prog Lipid Res, 2009. **48**(5): p. 298-305.
97. Drin, G. and B. Antonny, *Amphipathic helices and membrane curvature.* FEBS Lett, 2010. **584**(9): p. 1840-7.
98. Ford, M.G., et al., *Curvature of clathrin-coated pits driven by epsin.* Nature, 2002. **419**(6905): p. 361-6.
99. Bigay, J., et al., *ArfGAP1 responds to membrane curvature through the folding of a lipid packing sensor motif.* EMBO J, 2005. **24**(13): p. 2244-53.
100. Antonny, B., et al., *Membrane curvature and the control of GTP hydrolysis in Arf1 during COPI vesicle formation.* Biochem Soc Trans, 2005. **33**(Pt 4): p. 619-22.
101. Mesmin, B., et al., *Two lipid-packing sensor motifs contribute to the sensitivity of ArfGAP1 to membrane curvature.* Biochemistry, 2007. **46**(7): p. 1779-90.
102. Hurley, J.H., *Membrane binding domains.* Biochim Biophys Acta, 2006. **1761**(8): p. 805-11.
103. Lemmon, M.A. and K.M. Ferguson, *Signal-dependent membrane targeting by pleckstrin homology (PH) domains.* Biochem J, 2000. **350 Pt 1**: p. 1-18.
104. Lemmon, M.A., K.M. Ferguson, and J. Schlessinger, *PH domains: diverse sequences with a common fold recruit signaling molecules to the cell surface.* Cell, 1996. **85**(5): p. 621-4.
105. Lemmon, M.A., *Pleckstrin homology (PH) domains and phosphoinositides.* Biochem Soc Symp, 2007(74): p. 81-93.
106. Rebecchi, M.J. and S. Scarlata, *Pleckstrin homology domains: a common fold with diverse functions.* Annu Rev Biophys Biomol Struct, 1998. **27**: p. 503-28.
107. Lemmon, M.A., *Pleckstrin homology domains: not just for phosphoinositides.* Biochem Soc Trans, 2004. **32**(Pt 5): p. 707-11.
108. Ferguson, K.M., et al., *Structure of the high affinity complex of inositol trisphosphate with a phospholipase C pleckstrin homology domain.* Cell, 1995. **83**(6): p. 1037-46.
109. Thomas, C.C., et al., *Crystal structure of the phosphatidylinositol 3,4-bisphosphate-binding pleckstrin homology (PH) domain of tandem PH-domain-containing protein 1 (TAPP1): molecular basis of lipid specificity.* Biochem J, 2001. **358**(Pt 2): p. 287-94.

110. Ferguson, A.D., et al., *Crystal structure of the antibiotic albomycin in complex with the outer membrane transporter FhuA.* Protein Sci, 2000. **9**(5): p. 956-63.
111. Cronin, T.C., et al., *Structural determinants of phosphoinositide selectivity in splice variants of Grp1 family PH domains.* EMBO J, 2004. **23**(19): p. 3711-20.
112. Klarlund, J.K., et al., *Distinct polyphosphoinositide binding selectivities for pleckstrin homology domains of GRP1-like proteins based on diglycine versus triglycine motifs.* J Biol Chem, 2000. **275**(42): p. 32816-21.
113. Klarlund, J.K., et al., *Regulation of GRP1-catalyzed ADP ribosylation factor guanine nucleotide exchange by phosphatidylinositol 3,4,5-trisphosphate.* J Biol Chem, 1998. **273**(4): p. 1859-62.
114. Lietzke, S.E., et al., *Structural basis of 3-phosphoinositide recognition by pleckstrin homology domains.* Mol Cell, 2000. **6**(2): p. 385-94.
115. DiNitto, J.P., T.C. Cronin, and D.G. Lambright, *Membrane recognition and targeting by lipid-binding domains.* Sci STKE, 2003. **2003**(213): p. re16.
116. Carpten, J.D., et al., *A transforming mutation in the pleckstrin homology domain of AKT1 in cancer.* Nature, 2007. **448**(7152): p. 439-44.
117. Landgraf, K.E., C. Pilling, and J.J. Falke, *Molecular mechanism of an oncogenic mutation that alters membrane targeting: Glu17Lys modifies the PIP lipid specificity of the AKT1 PH domain.* Biochemistry, 2008. **47**(47): p. 12260-9.
118. Do, T.V., et al., *The role of activin A and Akt/GSK signaling in ovarian tumor biology.* Endocrinology, 2008. **149**(8): p. 3809-16.
119. Pilling, C., K.E. Landgraf, and J.J. Falke, *The GRP1 PH domain, like the AKT1 PH domain, possesses a sentry glutamate residue essential for specific targeting to plasma membrane PI(3,4,5)P(3).* Biochemistry, 2011. **50**(45): p. 9845-56.
120. Corbin, J.A., R.A. Dirkx, and J.J. Falke, *GRP1 pleckstrin homology domain: activation parameters and novel search mechanism for rare target lipid.* Biochemistry, 2004. **43**(51): p. 16161-73.
121. Knight, J.D. and J.J. Falke, *Single-molecule fluorescence studies of a PH domain: new insights into the membrane docking reaction.* Biophys J, 2009. **96**(2): p. 566-82.
122. Chen, H.C., et al., *Membrane docking geometry of GRP1 PH domain bound to a target lipid bilayer: an EPR site-directed spin-labeling and relaxation study.* PLoS One, 2012. **7**(3): p. e33640.
123. Lai, C.L., et al., *Molecular mechanism of membrane binding of the GRP1 PH domain.* J Mol Biol, 2013. **425**(17): p. 3073-90.
124. Lumb, C.N., et al., *Biophysical and computational studies of membrane penetration by the GRP1 pleckstrin homology domain.* Structure, 2011. **19**(9): p. 1338-46.
125. Ziemba, B.P. and J.J. Falke, *Lateral diffusion of peripheral membrane proteins on supported lipid bilayers is controlled by the additive frictional drags of (1) bound lipids and (2) protein domains penetrating into the bilayer hydrocarbon core.* Chem Phys Lipids, 2013. **172-173**: p. 67-77.

126. Ogasawara, M., et al., *Similarities in function and gene structure of cytohesin-4 and cytohesin-1, guanine nucleotide-exchange proteins for ADP-ribosylation factors.* J Biol Chem, 2000. **275**(5): p. 3221-30.
127. Fuss, B., et al., *The cytohesin Steppke is essential for insulin signalling in Drosophila.* Nature, 2006. **444**(7121): p. 945-8.
128. Hafner, M., et al., *Inhibition of cytohesins by SecinH3 leads to hepatic insulin resistance.* Nature, 2006. **444**(7121): p. 941-4.
129. Klarlund, J.K., et al., *Signaling complexes of the FERM domain-containing protein GRSP1 bound to ARF exchange factor GRP1.* J Biol Chem, 2001. **276**(43): p. 40065-70.
130. DiNitto, J.P., et al., *Specificity and membrane partitioning of Grsp1 signaling complexes with Grp1 family Arf exchange factors.* Biochemistry, 2010. **49**(29): p. 6083-92.
131. Mansour, M., S.Y. Lee, and B. Pohajdak, *The N-terminal coiled coil domain of the cytohesin/ARNO family of guanine nucleotide exchange factors interacts with the scaffolding protein CASP.* J Biol Chem, 2002. **277**(35): p. 32302-9.
132. Nevrivy, D.J., et al., *Interaction of GRASP, a protein encoded by a novel retinoic acid-induced gene, with members of the cytohesin family of guanine nucleotide exchange factors.* J Biol Chem, 2000. **275**(22): p. 16827-36.
133. DiNitto, J.P., et al., *Structural basis and mechanism of autoregulation in 3-phosphoinositide-dependent Grp1 family Arf GTPase exchange factors.* Mol Cell, 2007. **28**(4): p. 569-83.
134. Hiester, K.G. and L.C. Santy, *The cytohesin coiled-coil domain interacts with threonine 276 to control membrane association.* PLoS One, 2013. **8**(11): p. e82084.
135. Peterson, J.R. and E.A. Golemis, *Autoinhibited proteins as promising drug targets.* J Cell Biochem, 2004. **93**(1): p. 68-73.
136. Freedman, T.S., et al., *A Ras-induced conformational switch in the Ras activator Son of sevenless.* Proc Natl Acad Sci U S A, 2006. **103**(45): p. 16692-7.
137. Richardson, B.C. and J.C. Fromme, *Autoregulation of Sec7 Arf-GEF activity and localization by positive feedback.* Small GTPases, 2012. **3**(4): p. 240-3.
138. Sondermann, H., et al., *Structural analysis of autoinhibition in the Ras activator Son of sevenless.* Cell, 2004. **119**(3): p. 393-405.
139. He, X., et al., *Structural basis for autoinhibition of the guanine nucleotide exchange factor FARP2.* Structure, 2013. **21**(3): p. 355-64.
140. Iwig, J.S., et al., *Structural analysis of autoinhibition in the Ras-specific exchange factor RasGRP1.* Elife, 2013. **2**: p. e00813.
141. Gureasko, J., et al., *Role of the histone domain in the autoinhibition and activation of the Ras activator Son of Sevenless.* Proc Natl Acad Sci U S A, 2010. **107**(8): p. 3430-5.
142. Gureasko, J., et al., *Membrane-dependent signal integration by the Ras activator Son of sevenless.* Nat Struct Mol Biol, 2008. **15**(5): p. 452-61.

143. Aghazadeh, B., et al., *Structural basis for relief of autoinhibition of the Dbl homology domain of proto-oncogene Vav by tyrosine phosphorylation.* Cell, 2000. **102**(5): p. 625-33.
144. Li, P., et al., *Internal dynamics control activation and activity of the autoinhibited Vav DH domain.* Nat Struct Mol Biol, 2008. **15**(6): p. 613-8.
145. Yu, B., et al., *Structural and energetic mechanisms of cooperative autoinhibition and activation of Vav1.* Cell, 2010. **140**(2): p. 246-56.
146. Franzusoff, A., et al., *Localization of components involved in protein transport and processing through the yeast Golgi apparatus.* J Cell Biol, 1991. **112**(1): p. 27-37.
147. Richardson, B.C., C.M. McDonold, and J.C. Fromme, *The Sec7 Arf-GEF is recruited to the trans-Golgi network by positive feedback.* Dev Cell, 2012. **22**(4): p. 799-810.
148. Frank, S.R., J.C. Hatfield, and J.E. Casanova, *Remodeling of the actin cytoskeleton is coordinately regulated by protein kinase C and the ADP-ribosylation factor nucleotide exchange factor ARNO.* Mol Biol Cell, 1998. **9**(11): p. 3133-46.
149. Santy, L.C., et al., *Regulation of ARNO nucleotide exchange by a PH domain electrostatic switch.* Curr Biol, 1999. **9**(20): p. 1173-6.
150. Dierks, H., J. Kolanus, and W. Kolanus, *Actin cytoskeletal association of cytohesin-1 is regulated by specific phosphorylation of its carboxyl-terminal polybasic domain.* J Biol Chem, 2001. **276**(40): p. 37472-81.
151. Stalder, D., et al., *Kinetic studies of the Arf activator Arno on model membranes in the presence of Arf effectors suggest control by a positive feedback loop.* J Biol Chem, 2011. **286**(5): p. 3873-83.
152. Bos, J.L., H. Rehmann, and A. Wittinghofer, *GEFs and GAPs: critical elements in the control of small G proteins.* Cell, 2007. **129**(5): p. 865-77.
153. Lemmon, M.A., *Pleckstrin homology domains: not just for phosphoinositides.* Biochemical Society transactions, 2004. **32**(Pt 5): p. 707-11.
154. DiNitto, J.P. and D.G. Lambright, *Membrane and juxtamembrane targeting by PH and PTB domains.* Biochim Biophys Acta, 2006. **1761**(8): p. 850-67.
155. Donaldson, J.G. and C.L. Jackson, *ARF family G proteins and their regulators: roles in membrane transport, development and disease.* Nature reviews. Molecular cell biology, 2011. **12**(6): p. 362-75.
156. DiNitto, J.P., et al., *Structural basis and mechanism of autoregulation in 3-phosphoinositide-dependent Grp1 family Arf GTPase exchange factors.* Molecular cell, 2007. **28**(4): p. 569-83.
157. Gureasko, J., et al., *Membrane-dependent signal integration by the Ras activator Son of sevenless.* Nature structural & molecular biology, 2008. **15**(5): p. 452-61.
158. Gureasko, J., et al., *Role of the histone domain in the autoinhibition and activation of the Ras activator Son of Sevenless.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(8): p. 3430-5.
159. Nie, Z., D.S. Hirsch, and P.A. Randazzo, *Arf and its many interactors.* Current opinion in cell biology, 2003. **15**(4): p. 396-404.

160. Gillingham, A.K. and S. Munro, *The small G proteins of the Arf family and their regulators.* Annual review of cell and developmental biology, 2007. **23**: p. 579-611.
161. Franco, M., et al., *Myristoylation of ADP-ribosylation factor 1 facilitates nucleotide exchange at physiological Mg2+ levels.* The Journal of biological chemistry, 1995. **270**(3): p. 1337-41.
162. Randazzo, P.A., et al., *The myristoylated amino terminus of ADP-ribosylation factor 1 is a phospholipid- and GTP-sensitive switch.* The Journal of biological chemistry, 1995. **270**(24): p. 14809-15.
163. Liu, Y., R.A. Kahn, and J.H. Prestegard, *Dynamic structure of membrane-anchored Arf*GTP.* Nature structural & molecular biology, 2010. **17**(7): p. 876-81.
164. Li, J., et al., *Grp1 plays a key role in linking insulin signaling to glut4 recycling.* Developmental cell, 2012. **22**(6): p. 1286-98.
165. Kolanus, W., et al., *Alpha L beta 2 integrin/LFA-1 binding to ICAM-1 induced by cytohesin-1, a cytoplasmic regulatory molecule.* Cell, 1996. **86**(2): p. 233-42.
166. Klarlund, J.K., et al., *Signaling by phosphoinositide-3,4,5-trisphosphate through proteins containing pleckstrin and Sec7 homology domains.* Science, 1997. **275**(5308): p. 1927-30.
167. Rameh, L.E., et al., *A comparative analysis of the phosphoinositide binding specificity of pleckstrin homology domains.* J Biol Chem, 1997. **272**(35): p. 22059-66.
168. Kavran, J.M., et al., *Specificity and promiscuity in phosphoinositide binding by pleckstrin homology domains.* J Biol Chem, 1998. **273**(46): p. 30497-508.
169. Langille, S.E., et al., *ADP-ribosylation factor 6 as a target of guanine nucleotide exchange factor GRP1.* J Biol Chem, 1999. **274**(38): p. 27099-104.
170. Macia, E., S. Paris, and M. Chabre, *Binding of the PH and polybasic C-terminal domains of ARNO to phosphoinositides and to acidic lipids.* Biochemistry, 2000. **39**(19): p. 5893-901.
171. Klarlund, J.K., et al., *Distinct polyphosphoinositide binding selectivities for pleckstrin homology domains of GRP1-like proteins based on diglycine versus triglycine motifs.* J. Biol. Chem., 2000. **275**(42): p. 32816-21.
172. Venkateswarlu, K., et al., *Insulin-dependent translocation of ARNO to the plasma membrane of adipocytes requires phosphatidylinositol 3-kinase.* Curr Biol, 1998. **8**(8): p. 463-6.
173. Nagel, W., et al., *The PH domain and the polybasic c domain of cytohesin-1 cooperate specifically in plasma membrane association and cellular function.* Mol Biol Cell, 1998. **9**(8): p. 1981-94.
174. Venkateswarlu, K., et al., *EGF-and NGF-stimulated translocation of cytohesin-1 to the plasma membrane of PC12 cells requires PI 3-kinase activation and a functional cytohesin-1 PH domain.* J Cell Sci, 1999. **112 (Pt 12)**: p. 1957-65.
175. Cohen, L.A., et al., *Active Arf6 recruits ARNO/cytohesin GEFs to the PM by binding their PH domains.* Molecular biology of the cell, 2007. **18**(6): p. 2244-53.

176. Li, C.C., et al., *ARL4D recruits cytohesin-2/ARNO to modulate actin remodeling.* Molecular biology of the cell, 2007. **18**(11): p. 4420-37.
177. Stalder, D., et al., *Kinetic studies of the Arf activator Arno on model membranes in the presence of Arf effectors suggest control by a positive feedback loop.* The Journal of biological chemistry, 2011. **286**(5): p. 3873-83.
178. Ferguson, K.M., et al., *Structural basis for discrimination of 3-phosphoinositides by pleckstrin homology domains.* Molecular cell, 2000. **6**(2): p. 373-84.
179. Mossessova, E., R.A. Corpina, and J. Goldberg, *Crystal structure of ARF1*Sec7 complexed with Brefeldin A and its implications for the guanine nucleotide exchange mechanism.* Molecular cell, 2003. **12**(6): p. 1403-11.
180. He, J., et al., *Molecular mechanism of membrane targeting by the GRP1 PH domain.* Journal of lipid research, 2008. **49**(8): p. 1807-15.
181. Pasqualato, S., et al., *The structural GDP/GTP cycle of human Arf6.* EMBO reports, 2001. **2**(3): p. 234-8.
182. Menetrey, J., et al., *Structural basis for ARF1-mediated recruitment of ARHGAP21 to Golgi membranes.* The EMBO journal, 2007. **26**(7): p. 1953-62.
183. Varnai, P., et al., *Selective cellular effects of overexpressed pleckstrin-homology domains that recognize PtdIns(3,4,5)P3 suggest their interaction with protein binding partners.* Journal of cell science, 2005. **118**(Pt 20): p. 4879-88.
184. Isabet, T., et al., *The structural basis of Arf effector specificity: the crystal structure of ARF6 in a complex with JIP4.* The EMBO journal, 2009. **28**(18): p. 2835-45.
185. Cavenagh, M.M., et al., *Intracellular distribution of Arf proteins in mammalian cells. Arf6 is uniquely localized to the plasma membrane.* The Journal of biological chemistry, 1996. **271**(36): p. 21767-74.
186. Heller, H., M. Schaefer, and K. Schulten, *Molecular dynamics simulation of a bilayer of 200 lipids in the gel and in the liquid crystal phase.* The Journal of Physical Chemistry, 1993. **97**(31): p. 8343-8360.
187. Otwinowski, Z. and W. Minor, *Processing of X-ray Diffraction data Collected in Oscillation mode.* Methods Enzymol, 1997. **276**(part A): p. 307-326.
188. Langer, G., et al., *Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7.* Nat Protoc, 2008. **3**(7): p. 1171-9.
189. Murshudov, G.N., et al., *REFMAC5 for the refinement of macromolecular crystal structures.* Acta crystallographica. Section D, Biological crystallography, 2011. **67**(Pt 4): p. 355-67.
190. Cowtan, K., *The Buccaneer software for automated model building. 1. Tracing protein chains.* Acta crystallographica. Section D, Biological crystallography, 2006. **62**(Pt 9): p. 1002-11.
191. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
192. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.

193. Putnam, C.D., et al., *X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution.* Q Rev Biophys, 2007. **40**(3): p. 191-285.
194. Petoukhov, M.V. and D.I. Svergun, *Applications of small-angle X-ray scattering to biomacromolecular solutions.* Int J Biochem Cell Biol, 2013. **45**(2): p. 429-37.
195. Lipfert, J., et al., *Size and shape of detergent micelles determined by small-angle X-ray scattering.* J Phys Chem B, 2007. **111**(43): p. 12427-38.
196. Lipfert, J. and S. Doniach, *Small-angle X-ray scattering from RNA, proteins, and protein complexes.* Annu Rev Biophys Biomol Struct, 2007. **36**: p. 307-27.
197. Chen, B., et al., *Multiple conformations of SAM-II riboswitch detected with SAXS and NMR spectroscopy.* Nucleic Acids Res, 2012. **40**(7): p. 3117-30.
198. Davies, J.M., et al., *Conformational changes of p97 during nucleotide hydrolysis determined by small-angle X-Ray scattering.* Structure, 2005. **13**(2): p. 183-95.
199. Hammel, M., et al., *Structural basis of cellulosome efficiency explored by small angle X-ray scattering.* J Biol Chem, 2005. **280**(46): p. 38562-8.
200. Wang, J., et al., *A method for helical RNA global structure determination in solution using small-angle x-ray scattering and NMR measurements.* J Mol Biol, 2009. **393**(3): p. 717-34.
201. Mylonas, E. and D.I. Svergun, *Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering.* J Appl Crystallogr, 2007. **40**(s1): p. s245-s249.
202. Hammel, M., et al., *Solution structure of human and bovine beta(2)-glycoprotein I revealed by small-angle X-ray scattering.* J Mol Biol, 2002. **321**(1): p. 85-97.
203. Perez, J. and Y. Nishino, *Advances in X-ray scattering: from solution SAXS to achievements with coherent beams.* Curr Opin Struct Biol, 2012. **22**(5): p. 670-8.
204. David, G. and J. Perez, *Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline.* J Appl Crystallogr, 2009. **42**(5): p. 892-900.
205. Watanabe, Y. and Y. Inoko, *Size-exclusion chromatography combined with small-angle X-ray scattering optics.* J Chromatogr A, 2009. **1216**(44): p. 7461-5.
206. Gunn, N.J., et al., *Purification, crystallization, small-angle X-ray scattering and preliminary X-ray diffraction analysis of the SH2 domain of the Csk-homologous kinase.* Acta Crystallogr Sect F Struct Biol Cryst Commun, 2011. **67**(Pt 3): p. 336-9.
207. Chacon, P., et al., *Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm.* Biophys J, 1998. **74**(6): p. 2760-75.
208. Chacon, P., et al., *Reconstruction of protein form with X-ray solution scattering and a genetic algorithm.* J Mol Biol, 2000. **299**(5): p. 1289-302.
209. Heller, W.T., J.K. Krueger, and J. Trewhella, *Further insights into calmodulin-myosin light chain kinase interaction from solution scattering and shape restoration.* Biochemistry, 2003. **42**(36): p. 10579-88.

210. Heller, W., *ELLSTAT: shape modeling for solution small-angle scattering of proteins and protein complexes with automated statistical characterization.* J Appl Crystallogr, 2006. **39**(5): p. 671-675.
211. Zhao, J., et al., *Quaternary structures of a catalytic subunit-regulatory subunit dimeric complex and the holoenzyme of the cAMP-dependent protein kinase by neutron contrast variation.* J Biol Chem, 1998. **273**(46): p. 30448-59.
212. Walther, D., F.E. Cohen, and S. Doniach, *Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray solution scattering data for biomolecules.* J Appl Crystallogr, 2000. **33**(2): p. 350-363.
213. Konarev, P.V., et al., *ATSAS 2.1, a program package for small-angle scattering data analysis.* J Appl Crystallogr, 2006. **39**(2): p. 277-286.
214. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophys J, 1999. **76**(6): p. 2879-86.
215. Franke, D. and D.I. Svergun, *DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering.* J Appl Crystallogr, 2009. **42**(2): p. 342-346.
216. Svergun, D.I., M.V. Petoukhov, and M.H. Koch, *Determination of domain structure of proteins from X-ray solution scattering.* Biophys J, 2001. **80**(6): p. 2946-53.
217. Volkov, V.V. and D.I. Svergun, *Uniqueness of ab initio shape determination in small-angle scattering.* J Appl Crystallogr, 2003. **36**(3 Part 1): p. 860-864.
218. Lambright, D.G., et al., *Complementary techniqes enance the quality and scope of information obtained from SAXS.* Transactions of the American Crystallography Society, 2013. **44**: p. 1-12.
219. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* J Appl Crystallogr, 2003. **36**(5): p. 1277-1282.
220. Fetler, L., et al., *X-ray scattering titration of the quaternary structure transition of aspartate transcarbamylase with a bisubstrate analogue: influence of nucleotide effectors.* J Mol Biol, 1995. **251**(2): p. 243-55.
221. Perez, J., et al., *Heat-induced unfolding of neocarzinostatin, a small all-beta protein investigated by small-angle X-ray scattering.* J Mol Biol, 2001. **308**(4): p. 721-43.
222. Savitzky, A. and M.J.E. Golay, *Smoothing and Differentiation of Data by Simplified Least Squares Procedures.* Anal Chem, 1964. **36**(8): p. 1627-1639.
223. Kratky, O. and G. Porod, *Diffuse small-angle scattering of X-rays in colloid systems.* J Colloid Sci, 1949. **4**(1): p. 35-70.
224. Porod, G., *X-ray small angle scattering of close packed colloidal systems.* Kolloid Zeitschrift, 1951. **124**: p. 83-114.
225. Porod, G., *"General Theory" in Small Angle X-ray Scattering*. O. Glatter and O. Kratky ed1982, London: Academic Press.
226. Harpaz, Y., M. Gerstein, and C. Chothia, *Volume changes on protein folding.* Structure, 1994. **2**(7): p. 641-9.

227. Guinier, A. and G.r. Fournet, *Small-angle scattering of X-rays*. Structure of matter series1955, New York,: Wiley. 268 p.
228. Glatter, O., *A new method for the evaluation of small-angle scattering data.* J Appl Crystallogr, 1977. **10**(5): p. 415-421.
229. Glatter, O., *The interpretation of real-space information from small-angle scattering experiments.* J Appl Crystallogr, 1979. **12**(2): p. 166-175.
230. Semenyuk, A.V. and D.I. Svergun, *GNOM - a program package for small-angle scattering data processing.* J Appl Crystallogr, 1991. **24**(5): p. 537-540.
231. Rosenberg, O.S., et al., *Structure of the autoinhibited kinase domain of CaMKII and SAXS analysis of the holoenzyme.* Cell, 2005. **123**(5): p. 849-60.
232. Kozin, M.B. and D.I. Svergun, *Automated matching of high- and low-resolution structural models.* J Appl Crystallogr, 2001. **34**(1): p. 33-41.
233. Alix, E., et al., *The capping domain in RalF regulates effector functions.* PLoS Pathog, 2012. **8**(11): p. e1003012.
234. Malaby, A.W., B. van den Berg, and D.G. Lambright, *Structural basis for membrane recruitment and allosteric activation of cytohesin family Arf GTPase exchange factors.* Proc Natl Acad Sci U S A, 2013. **110**(35): p. 14213-8.
235. Nie, Z., D.S. Hirsch, and P.A. Randazzo, *Arf and its many interactors.* Curr Opin Cell Biol, 2003. **15**(4): p. 396-404.
236. Franco, M., et al., *Myristoylation of ADP-ribosylation factor 1 facilitates nucleotide exchange at physiological Mg2+ levels.* J Biol Chem, 1995. **270**(3): p. 1337-41.
237. Randazzo, P.A., et al., *The myristoylated amino terminus of ADP-ribosylation factor 1 is a phospholipid- and GTP-sensitive switch.* J Biol Chem, 1995. **270**(24): p. 14809-15.
238. Svergun, D.I., et al., *New Developments in Direct Shape Determination from Small-Angle Scattering. 2. Uniqueness.* Acta Crystallographica Section A, 1996. **52**(3): p. 419-426.
239. Peters, P.J., et al., *Overexpression of wild-type and mutant ARF1 and ARF6: distinct perturbations of nonoverlapping membrane compartments.* J Cell Biol, 1995. **128**(6): p. 1003-17.
240. Bernado, P., et al., *Structural characterization of flexible proteins using small-angle X-ray scattering.* J Am Chem Soc, 2007. **129**(17): p. 5656-64.
241. Petoukhov, M.V. and D.I. Svergun, *Global rigid body modeling of macromolecular complexes against small-angle scattering data.* Biophys J, 2005. **89**(2): p. 1237-50.
242. Petoukhov, M.V., et al., *New developments in the ATSAS program package for small-angle scattering data analysis.* J Appl Crystallogr, 2012. **45**(2): p. 342-350.
243. Varnai, P., et al., *Selective cellular effects of overexpressed pleckstrin-homology domains that recognize PtdIns(3,4,5)P3 suggest their interaction with protein binding partners.* J Cell Sci, 2005. **118**(Pt 20): p. 4879-88.
244. Miao, B., et al., *Small molecule inhibition of phosphatidylinositol-3,4,5-triphosphate (PIP3) binding to pleckstrin homology domains.* Proc Natl Acad Sci U S A, 2010. **107**(46): p. 20126-31.

245. Miao, B., et al., *Inhibition of cell migration by PITENINs: the role of ARF6.* Oncogene, 2012. **31**(39): p. 4317-32.
246. Schweitzer, J.K. and C. D'Souza-Schorey, *Localization and activation of the ARF6 GTPase during cleavage furrow ingression and cytokinesis.* J Biol Chem, 2002. **277**(30): p. 27210-6.
247. Ismail, S.A., et al., *Arl2-GTP and Arl3-GTP regulate a GDI-like transport system for farnesylated cargo.* Nat Chem Biol, 2011. **7**(12): p. 942-9.
248. Ismail, S.A., et al., *Structural basis for Arl3-specific release of myristoylated ciliary cargo from UNC119.* EMBO J, 2012. **31**(20): p. 4085-94.
249. Folly-Klan, M., et al., *A novel membrane sensor controls the localization and ArfGEF activity of bacterial RalF.* PLoS Pathog, 2013. **9**(11): p. e1003747.
250. Del Campo, C.M., et al., *Structural Basis for PI(4)P-Specific Membrane Recruitment of the Legionella pneumophila Effector DrrA/SidM.* Structure, 2014. **22**(3): p. 397-408.
251. Pan, T., et al., *Cytohesins/ARNO: The Function in Colorectal Cancer Cells.* PLoS One, 2014. **9**(3): p. e90997.
252. Pan, T., et al., *Function and mode of action of cytohesins in the epidermal growth factor pathway in colorectal cancer cells.* Oncol Lett, 2013. **5**(2): p. 521-526.